Annual Price and Real Expenditure Comparisons within the International Comparison Program: Feasibility and Empirical Illustration

Robert Inklaar*
University of Groningen, Groningen, Netherlands– e-mail address

D.S. Prasada Rao
The University of Queensland, Brisbane, Australia – d.rao@uq.edu.au

Abstract

The current state of play with the International Comparison Program is that Purchasing Power Parity (PPP) exchange rates for currencies and internationally comparable national accounts aggregates are available roughly every five years at the global level. These are available slightly more frequently for the OECD and European Countries. The demand for international comparisons are currently largely met by extrapolations made available through the *World Development Indicators; Penn World Table,* and to a lesser extent, from sources such as the *Maddison Series* and series from *UQICD*. The extrapolation methodologies differ across sources and usually undertaken at an aggregate level. The need for reliable PPP and international comparisons data on annual bases is ever growing and the World Bank, which has overseen the implementation of ICP at the global level, is taking a lead role in exploring the feasibility of annual comparisons of PPPs at the regional and global level.

The paper explores various extrapolation strategies and use an analytical framework to identify the level of disaggregation necessary to produce reliable PPPs and international comparisons. The concept of a reduced information approach to international comparisons is described and then identifies the sources and methods to be used in assembling the basic data necessary for PPP extrapolation. In particular, we canvass the full use of all the information available from different sources in this process and then identify remaining gaps in data which may then need to be filled using targeted surveys or information from non-conventional sources. The paper demonstrates how the currently available information can be used in implementing ICP for the years 2012 and 2013 and provide meaningful annualized comparison following the 2011 benchmark comparison. Preliminary results of updated global PPPs are presented and compared with the extrapolations based on the current WDI updating approach.

Keywords: Purchasing power parities; extrapolation; international comparison program

1. Introduction

The International Comparison Program (ICP) is a major global statistical initiative undertaken under the auspices of the Statistical Commission of the United Nations. The main purpose of the ICP is to produce reliable and timely estimates of purchasing power parity (PPP) exchange rates and internationally comparable economic aggregates such as the gross domestic product (GDP), household consumption, government expenditure and gross fixed capital formation. The ICP which started as a small collaborative research project between researchers at the University of Pennsylvania and the UN Statistical Office has grown into a major global statistical initiative. The recently completed round of the ICP with 2011 as the benchmark year covered 199 countries including small island economies in the Pacific and the Caribbean, only 177 of these countries participated in comparisons at the full GDP level. A complete description of the methodology and the full set of results are available in World Bank (2015) as well as the dedicated ICP website located on the World Bank website.

The international community of researchers and policy makers has enthusiastically received results of ICP from the 2005 and 2011 ICP rounds which have covered, respectively, 146 and 177 countries at the GDP level. The World Bank continues to use ICP estimates of PPPs for household consumption as the basis for the calibration of international poverty line as well as for the estimation of incidence of regional and global poverty. The poverty line was set at USD 1.25 in 2005 and after much deliberation and background research the Bank has decided to set the international poverty line for 2011 using results from ICP 2011 and the line is now set at USD 1.85. The International Monetary Fund uses PPPs for internal governance purposes. This follows a long tradition of such use in the European Community. Eurostat currently produces estimates of PPPs and real expenditure comparisons on an annual basis.

Currently, PPP exchange rates for currencies and internationally comparable national accounts aggregates from ICP are available roughly every five years at the global level. These are available slightly more frequently for the OECD and EC countries. The demand for international comparisons are currently largely met by extrapolations made available through the World Development Indicators; Penn World Table, and to a lesser extent, from sources such as the Maddison Series and series from UQICD. The extrapolation methodologies differ across sources and usually undertaken at an aggregate level. The need for reliable PPP and international comparisons data on annual bases is ever growing and the World Bank, which has overseen the implementation of ICP at the global level, is taking a lead role in exploring the feasibility of annual comparisons of PPPs at the regional and global level.

The need for frequent international comparisons of prices and real expenditure is reinforced by the significant discrepancies between the benchmark comparisons observed at the time of release of the last two ICP comparisons in 2005 and 2011. When the 2005 results were released, the size of the global economy in 2005 was found to be some 40% smaller than the size indicated by extrapolations from the 1996 set of international comparisons. The global poverty and inequality were found to be much higher than those derived from the extrapolations. Differences of similar magnitude, but in the reverse direction, became apparent after the release of comparisons from the 2011 round of the ICP. These differences have prompted researchers to investigate the sources of differences. Deaton and Aten (2016) and Inklaar and Rao (2016) investigated the sources of differences between the 2005 and 2011 ICP rounds whereas Deaton and Heston (2011), Feenstra, Ma, Neary and Rao (2013) examined the sources of differences between the 2005 round of the ICP and extrapolations from 1996. While the sources of divergence between extrapolations and the benchmark comparisons have been investigated before (see McCarthy, 2013a and Timmer and Inklaar, 2013a), the significant and large observed difference between benchmarks and extrapolation is often attributed to the infrequency of conducting ICP benchmarks. Inklaar and Rao (2016) conclude their work by observing "We would also argue that our results are a case for more frequent price comparisons. Since a price comparison in any year will provide results that are measured with a sizeable degree of error, more frequent measurement will prevent outliers in any one year from distorting international comparisons of living standards for a substantial period of time. Furthermore, the introduction of new methodologies can more easily be phased in gradually if there is an ongoing statistical infrastructure for surveying prices, again leading to fewer major revisions."

Since the completion of the ICP in 2011, the program has undergone a thorough evaluation conducted by the UN Statistical Commission. Based on an evaluation of the ICP, The Statistical Commission at its 47th meeting has endorsed the continuation of ICP and recommended that the ICP is conducted on a more regular basis using the rolling price survey approach. Recognizing the need for reliable and frequent international comparisons of PPPs and real expenditures, the DECDG is keen to establish a framework for compiling PPPs and real expenditure on an annual basis. The financial and human resource implications and the enormity of the task of getting all the participating countries to devote sufficient resources are major barriers to implementing ICP on an annual basis. A possible option is to follow the Eurostat practice of using rolling price surveys in compiling annual PPPs.

The main objective of this paper is to examine various options available for this purpose and propose a feasible methodology for producing ICP results on an annual basis. The paper will present a blueprint for achieving the goal of annual compilation of PPPs by the year 2020.

The paper is organized as follows. In section 2, we discuss three extrapolation strategies and make a case for extrapolation at the most disaggregated level. In section 3, we use the index number framework to identify the optimum level of disaggregation that minimizes inconsistencies between benchmarks and extrapolated PPPs. Section 4 explores the notion of a reduced information approach to international comparisons and then identifies the sources and methods to be used in assembling the basic data necessary for PPP extrapolation. In particular, we canvass the full use of all the information available from different sources in this process and then identify remaining gaps in data which may then need to be filled using targeted surveys or information from non-conventional sources. Section 5 demonstrates how the currently available information can be used in implementing ICP for the years 2012 and 2013 and provide meaningful annualized comparison following the 2011 benchmark comparison. Section 6 presents results of updated global PPPs for 2012 and 2013 and compares our extrapolations at the component level with the existing updating approach at the GDP level. In Section 7 we present a feasible framework for the implementation of the recommendations from its 47th annual meeting. We provide a blueprint for the 2017 benchmark followed by annual international comparisons for the years 2018 to 2020. It is expected that by 2020 the World Bank will have an established program of compilation of PPPs and real expenditures on an annual basis.

Section 2 : Compilation of annual PPPs – feasible approaches

2.1 Conducting full-scale ICP every year

The objective of compiling annual PPPs and real aggregates is a fairly simple task if there were no constraints on resources available for the exercise. In this unlikely scenario, the participating counties, Regional Coordinating Agencies and the Global Office[1] simply replicate the tasks involved in the benchmark comparisons, say the 2011 ICP, leading to a new set of PPPs for the new benchmark year. In terms of data requirements, the participating countries provide: (i) prices from the regional product lists; (iii) prices for items in the global core list; (iii) expenditure weights for the 155 basic headings drawn from national accounts; and these data are then supplemented by productivity adjustment factors compiled for this purpose by the World Bank or drawn from some other source.[2] These data are then aggregated at various levels ultimately leading to price and real expenditure comparisons at the regional level as well as global level; and at the level of basic headings as well as for higher level aggregates.

2.2 Aggregate level extrapolation

This is a commonly used approach where PPPs are updated from one year to the other using the national income deflators. At the GDP level, the updating procedure is given by:

$$PPP_j^{t+1} = PPP_j^t \times \frac{GDPDef_j^{t,t+1}}{GDPDef_R^{t,t+1}}$$

where *R* represents the reference country. This updating procedure can be used at the GDP level or at the level of a chosen aggregate. The *World Development Indicators* uses this approach to update PPPs from benchmark year to other years. The Penn World Table uses this at the level of household

---

[1] See World Bank (2015) and McCarthy (2013b) for a description of the governance structure for the ICP.
[2] In ICP 2011, productivity adjustment factors were supplied by Timmer and Inklaar (2013b).

consumption, government expenditure and investment. The UQICD has been using extrapolation at the GDP level until the recent work on extrapolation of PPPs at the component level.[3]

Extrapolation at the aggregate level using national level deflators usually results in seriously inconsistent series of extrapolations and the benchmarks. This issue has been well researched (see Dalgaard and Sorensen, 2002; Deaton, 2012; McCarthy, 2013a; and Inklaar and Timmer 2013a). There is a general consensus that the differences arise as the concepts and methods used for temporal and spatial comparisons differ. We return to this aspect below. Changes in methods and revisions to data; differences in survey methods; and measurement errors in the compilation of PPPs and in national accounts can contribute to the observed inconsistencies between benchmarks and extrapolations.

2.3 The rolling price-survey approach

The task of replicating ICP in full each year is prohibitively costly; and participating countries as well as regional coordinating agencies and the World Bank are not in a position to allocate the resources necessary to implement ICP every year.[4] Recognizing this reality, the Eurostat has been using rolling price survey approach since 1990. For any given year, the rolling price survey approach relies on extrapolation of two-thirds of price data from the previous year using appropriate deflators and the remaining one-third of prices are collected in the year of compilation. According to this approach, Eurostat does not conduct a fully-fledged ICP-type of exercise in any of the years. However, in a cycle of three years all the products will have been priced in at least one of the years.

Implementation of a rolling price-survey approach requires a choice of appropriate deflators and the optimal level at which extrapolation is made for two-thirds of the price data. We comment on the rolling price-survey approach in Section 4 below.

Section 3: An analytical framework for identifying optimal level of disaggregation for extrapolation

We draw from a short but excellent exposition of the problem of updating by Deaton (2012). Though Deaton (2012) focuses on updating PPP exchange rates for consumption using consumer price indexes, it appears that his analysis is general enough to provide guidance in the choice of the level of disaggregation at which extrapolation is made using an appropriate price index number.

We consider the simple case of two countries where PPP is computed using Törnqvist index numbers. For simplicity, we assume that the same set of commodities enter PPP and national level index number computation. We also assume that the expenditure shares of commodities differ across countries but remain the same over time periods t and $t + 1$. Let p_{ij}^s represent the price of the ith commodity $(i = 1, 2, ..., N)$ in country $j (= 1, 2)$ in period s $(s = t$ or $t + 1)$. Let s_{ij} represent expenditure shares associated with commodity i in country $j (j = 1, 2)$.[5] We further let PPP_2^s represent purchasing power parity of currency of country 2 with country 1 as the reference country in period s.[6] Let P_j represent the price index in country j (1 and 2) over time t to $t + 1$. Then the logarithmic form of the three Törnqvist indices are given by:

---

[3] The UQICD approach differs from both the WDI and PWT in that it makes use of all the available benchmark information instead of using only one or two benchmarks.

[4] Even when ICP is conducted once in five years, participating countries find it difficult to devote sufficient national resources to conduct ICP specific price surveys. Until such time the ICP activity is integrated into the work plans of national statistical agencies, it would be difficult for the participating countries to collect and provide data necessary for compiling reliable PPPs and real expenditures.

[5] We do not have time superscript with expenditure share as we assume that expenditure shares remain the same over time. Expenditure shares tend to move slowly over time, so this is not a tenuous assumption.

[6] We drop subscript 1 with PPP for ease of notation.

$$\ln PPP_2^s = \frac{1}{2}\sum_{i=1}^{N}(s_{i1} + s_{i2})(\ln p_{i2}^s - \ln p_{i1}^s) \text{ for } s = t \text{ or } t+1 \tag{1}$$

$$\ln P_2 = \sum_{i=1}^{N} s_{i2}(\ln p_{i2}^{t+1} - \ln p_{i2}^t) \tag{2}$$

$$\ln P_1 = \sum_{i=1}^{N} s_{i1}(\ln p_{i1}^{t+1} - \ln p_{i1}^t) \tag{3}$$

It is easy to see that PPP_2^s is a Törnqvist index that compares price levels across countries 1 and 2 whereas P_1 and P_2 represent Törnqvist indices for countries 1 and 2 measuring price changes from t to $t+1$.

Following Deaton (2012), we consider the change in PPP over time in logarithmic form. This is given by:

$$\ln PPP_2^{t+1} - \ln PPP_2^t = \frac{1}{2}\sum_{i=1}^{N}(s_{i1} + s_{i2})[(\ln p_{i2}^{t+1} - \ln p_{i1}^{t+1}) - (\ln p_{i2}^t - \ln p_{i1}^t)] \tag{4}$$

After simple rearrangement and definitions in (1), (2) and (3), we can show that equation (4) equals:

$$\ln PPP_2^{t+1} - \ln PPP_2^t = \ln P_2 - \ln P_1 - \frac{1}{2}\sum_{i=1}^{N}(s_{i2} - s_{i1})\left[\ln\left(\frac{p_{i2}^{t+1}}{p_{i2}^t}\right) + \ln\left(\frac{p_{i1}^{t+1}}{p_{i1}^t}\right)\right] \tag{5}$$

From equation (5), inconsistency between benchmark and updates is given by:

$$\ln PPP_2^{t+1} - \ln PPP_2^t - (\ln P_2 - \ln P_1) = -\frac{1}{2}\sum_{i=1}^{N}(s_{i2} - s_{i1})\left[\ln\left(\frac{p_{i2}^{t+1}}{p_{i2}^t}\right) + \ln\left(\frac{p_{i1}^{t+1}}{p_{i1}^t}\right)\right] \tag{6}$$

Deaton (2012) argues that this inconsistency depends on the covariance between differences in expenditure shares in the two countries and price movements in prices in the two countries under consideration.

However, we consider a different angle for equation (6). If the N commodities considered here represent a commodity group, we ask the question as to when the inconsistency between updates and benchmarks is likely to zero or very small. The following result provides a useful direction.

Result 1: Under the set-up considered in equations (1) to (6) based on Törnqvist index for the measurement of price levels across countries and price change over time, inconsistency between benchmarks and updating using domestic measures of price changes vanishes if all the commodities considered in the computation show the same price change over time.

In order to verify this result, suppose prices of all the commodities in country 2 change by the same percentage α and price change is uniform across commodities in country 1 represented by a percentage change β, then equation (6) becomes:

$$\ln PPP_2^{t+1} - \ln PPP_2^t - (\ln P_2 - \ln P_1) = -\frac{1}{2}\sum_{i=1}^{N}(s_{i2} - s_{i1})[\alpha + \beta]$$

$$= -\frac{1}{2}(\alpha + \beta)\sum_{i=1}^{N}(s_{i2} - s_{i1}) = 0 \tag{7}$$

The last equality in equation (7) follows from the fact that expenditure shares add up to 1.

We observe that the result reported here is based on the Törnqvist index. However it is easy to show that this result holds even when other index number formulae are used. Two further results are stated and proved below.

Result 2: Under the set-up considered in equations (1) to (6) and if the Fisher index is used for the purpose of price comparisons across countries and over time then the inconsistency between benchmarks and updating using domestic measures of price changes vanishes if all the commodities considered in the computation exhibit the same price change over time.

Given that all commodities exhibit the same level of price change over time, we can write the prices in period 2 for countries 1 and 2 respectively as:

$$p_{i2}^{t+1} = \beta \cdot p_{i2}^{t} \qquad and \qquad p_{i1}^{t+1} = \alpha \cdot p_{i}^{t} \tag{8}$$

Consider the Fisher index which is the geometric mean of the Laspeyres and Paasche indices. Given (8) it follows that the price change from period t to $t+1$ for counties 1 and 2 are respectively α and β, that is $P_2 = \beta \quad and \quad P_1 = \alpha$.

Now we consider the change in the price level for country 2 with country 1 as the reference country. This is given by the ratio:

$$\frac{P_2^{t+1}}{P_2^{t}} = \frac{\left[\dfrac{\sum_{i=1}^{N} p_{i2}^{t+1} q_{i2}^{t+1}}{\sum_{i=1}^{N} p_{i1}^{t+1} q_{i2}^{t+1}} \cdot \dfrac{\sum_{i=1}^{N} p_{i2}^{t+1} q_{i1}^{t+1}}{\sum_{i=1}^{N} p_{i1}^{t+1} q_{i1}^{t+1}}\right]^{0.5}}{\left[\dfrac{\sum_{i=1}^{N} p_{i2}^{t} q_{i2}^{t}}{\sum_{i=1}^{N} p_{i1}^{t} q_{i2}^{t}} \cdot \dfrac{\sum_{i=1}^{N} p_{i2}^{t} q_{i1}^{t}}{\sum_{i=1}^{N} p_{i1}^{t} q_{i1}^{t}}\right]^{0.5}} \tag{9}$$

Substituting (8) into (9) and observing that the expenditure shares remain the same over time, we can show after simple algebraic manipulations that

$$\frac{P_2^{t+1}}{P_2^{t}} = \frac{\left[\dfrac{\sum_{i=1}^{N} p_{i2}^{t+1} q_{i2}^{t+1}}{\sum_{i=1}^{N} p_{i1}^{t+1} q_{i2}^{t+1}} \cdot \dfrac{\sum_{i=1}^{N} p_{i2}^{t+1} q_{i1}^{t+1}}{\sum_{i=1}^{N} p_{i1}^{t+1} q_{i1}^{t+1}}\right]^{0.5}}{\left[\dfrac{\sum_{i=1}^{N} p_{i2}^{t} q_{i2}^{t}}{\sum_{i=1}^{N} p_{i1}^{t} q_{i2}^{t}} \cdot \dfrac{\sum_{i=1}^{N} p_{i2}^{t} q_{i1}^{t}}{\sum_{i=1}^{N} p_{i1}^{t} q_{i1}^{t}}\right]^{0.5}} = \frac{\left[\dfrac{\sum_{i=1}^{N} \beta p_{i2}^{t} q_{i2}^{t+1}}{\sum_{i=1}^{N} \alpha p_{i1}^{t} q_{i2}^{t+1}} \cdot \dfrac{\sum_{i=1}^{N} \beta p_{i2}^{t} q_{i1}^{t+1}}{\sum_{i=1}^{N} \alpha p_{i1}^{t} q_{i1}^{t+1}}\right]^{0.5}}{\left[\dfrac{\sum_{i=1}^{N} p_{i2}^{t} q_{i2}^{t}}{\sum_{i=1}^{N} p_{i1}^{t} q_{i2}^{t}} \cdot \dfrac{\sum_{i=1}^{N} p_{i2}^{t} q_{i1}^{t}}{\sum_{i=1}^{N} p_{i1}^{t} q_{i1}^{t}}\right]^{0.5}} = \frac{\beta}{\alpha} = \frac{P_2}{P_1} \tag{10}$$

Equation (10) implies: $\left(\ln P_2^{t+1} - \ln P_2^{t}\right) - \left(\ln P_2 - \ln P_1\right) = 0$, which in turn implies that there is no inconsistency between the benchmark comparisons and temporal price changes observed in countries 1 and 2.

Now we turn to a more general result that does not depend upon the functional form for the price index. Here a binary index that compares prices in period or country 2 with the base period or reference country 1, denoted by P_{12}, is a function of observed prices and quantities (p_2, p_1, q_1, q_2). We assume that the price index satisfies the following proportionality axioms[7]. The price index is given by a function of prices and quantities observed in the two periods/countries:

$$P_{12} = P(p_2, p_1, q_2, q_1) \tag{11}$$

Axiom of Proportionality in prices of current period: The price index P_{21} is said to satisfy this axiom if prices in period 2 are multiplied by a constant λ (> 0) then the index is itself multiplied by λ. That is:

$$P(\lambda p_2, p_1, q_1, q_2) = \lambda P(p_2, p_1, q_1, q_2) \tag{12}$$

---

[7] The axiomatic approach to index numbers is well researched. Comprehensive expositions of the axiomatic approach can be found in Balk (2008) and in ECE-ILO (2010) *Manual on the Consumer Price Index.*

Axiom of Proportionality in prices of base period: The price index P_{21} is said to satisfy this axiom if prices in period 1 are multiplied by a constant λ (>0) then the index is itself multiplied by $1/\lambda$. That is:

$$P(p_2, \lambda p_1, q_1, q_2) = \frac{1}{\lambda} P(p_2, p_1, q_1, q_2) \qquad (13)$$

The following result provides a sufficient condition for the consistency between benchmarks and temporal price movements.

Result 3: If the price index formula used for comparisons of prices across countries and over time are represented by a generic price index formula $P_{12} = P(p_2, p_1, q_2, q_1)$ and if the index satisfies the axioms of proportionality in current and base period/country prices, then the cross-country price comparisons across two different benchmarks are consistent with relative price movements in the two periods.

The proof follows from the definitions that use notation in equations (1) to (6). We have:

$$P_2^{t+1} = P(p_2^{t+1}, p_1^{t+1}, q_2^{t+1}, q_1^{t+1})$$
$$P_2^{t} = P(p_2^{t}, p_1^{t}, q_2^{t}, q_1^{t})$$
$$\mathbf{P}_2 = P(p_2^{t+1}, p_2^{t}, q_2^{t+1}, q_2^{t}) \qquad (14)$$
$$\mathbf{P}_1 = P(p_1^{t+1}, p_1^{t}, q_1^{t+1}, q_1^{t})$$

Making using of the fact that $p_2^{t+1} = \beta\, p_2^{t}$ *and* $p_1^{t+1} = \alpha\, p_1^{t}$ and using the two axioms stated above, we can show that

$$\frac{P_2^{t+1}}{P_2^{t}} = \frac{\mathbf{P}_2}{\mathbf{P}_1} = \frac{\beta}{\alpha} \qquad (15)$$

Therefore consistency between benchmarks and temporal price movements can be guaranteed in the case where price movements in the countries 1 and 2 are proportional and the index number formula used satisfies the two axioms of proportionality.
We make three observations:

1. The results stated here provide a sufficient condition but it is not a necessary condition. Further, the result is derived in a very special case.
2. We believe that this sufficient condition provides guidance as to the level of disaggregation at which we could extrapolate with minimum inconsistency. The answer according to the result is that the commodity group should be sufficiently homogeneous to exhibit similar price movements over time. In price index compilation, this concept is somewhat similar to commodity groups that underpin elementary indices.
3. This result suggests that it is best if extrapolation is undertaken at the basic heading level. It is generally expected that the products included in a basic heading are not only homogeneous but they also exhibit similar price level differences across countries and movements over time.

Adopting the framework considered in Deaton (2012) for our purpose of determining the optimum level of disaggregation, we find that extrapolation using national price deflators is best undertaken at the basic heading level. However, in actual implementation it may not be possible to obtain price deflators at a level of aggregation that matches the basic headings within the ICP. For example, the ICP has 110 consumption basic headings and most consumer price indices are available for 10 or 12 aggregate groups.

Section 4: Reduced information approach to annual compilation of PPPs

The costs associated with the implementation of a fully-fledged ICP every year are prohibitively high and therefore is not an alternative that can be considered. Recognizing this reality, it is necessary to identify a reduced data set that is likely to approximate PPPs from detailed data set.

PPP compilation requires two sets of data: prices and weights. Prices are collected through price surveys conducted in each of the participating countries whereas weights are typically based on expenditure weights from national accounts. Price surveys are carefully designed to suit the characteristics of the aggregate under consideration. Different approaches are used for collecting prices for consumption, government expenditure, construction, and machinery and equipment.

4.1 National accounts weights

National accounts expenditure weights are not usually available in real time. Further, national accounts data are subject to significant revisions. For the purpose of PPP computation, use of an average of weights of the three years prior to a given year may provide a stable set of weights. While NA aggregates tend to be revised, it is rare the weights are revised.

4.2 Household Consumption

Use of rolling price survey approach can play a significant role in providing reliable price data for household consumption PPPs. However, we believe that some discretion needs to be used in the implementation of the rolling price survey data.

1. It is important to divide the basic headings into two classes, one class of basic headings where items included do not exhibit major changes from year to year. For example, basic headings like meat, milk, fresh vegetables, etc. are basic headings where the items remain the same over time. On the other hand, basic headings that cover communications, computers, sport and recreation exhibit considerable change from year to year. For basic headings where products change in quality and also specifications it may be necessary to price them on an annual basis instead of relying on rolling price surveys.
2. From the framework discussed in Section 3, it is clear that price deflators used in extrapolation must closely match the basic headings.
3. Given that both 2005 and 2011 benchmarks were conducted using structured product descriptions (SPDs), it may be feasible to compute a basic heading level inflation figure using prices of products that closely match between the two benchmarks. Closeness of the ICP based basic heading inflation figures with the CPI based deflators can be used in determining the suitability of CPI deflators under consideration. In fact, this type of analysis was conducted by the Asian Development Bank (ADB, 2015) as a data editing procedure to identify possible outliers.
4. There are non-conventional sources of price data available, such as scanner data and internet prices, for the purpose of PPP compilation. These sources could be used to supplement the extrapolation procedure using national price deflators. These sources are likely to provide a more accurate estimate of price level differences when products and services are rapidly changing from year to year.
5. There are some items like electricity and water charges, postal rates and rates fixed by the government where it would be feasible to collect these prices on an annual basis.

4.3 Government Consumption

Compensation: Data on government compensation would be straightforward to obtain from government sources where wage increases are formally recorded. Even if wages and salaries by occupational classification may not be available, percentage changes in government salaries can be obtained from annual report of government departments.

Productivity adjustment: In ICP 2011 productivity adjustments were made to government compensation. The adjustment was based on productivity factors estimated by Inklaar and Timmer (2013b). Given relative productivity changes are gradual and also small from one year to the other, these factors may either be kept constant or a three-year moving average of productivity measures are used so that it eliminates noise in the estimation of productivity levels. In practical terms, this suggests that extrapolating from productivity-adjusted PPPs until a subsequent round of wage surveys is conducted would be a sensible approach.

Government consumption of goods and services: These may be dealt similar to household consumption where rolling survey approach is used for items that tend to remain stable over time and use annually collected price data for items exhibiting significant change from one year to the other.

4.4 Gross Capital Formation

The two components of Construction and Machinery and Equipment need to be dealt with separately. The current approach to Construction is easy to deal with as only prices of basic materials used in construction are needed. In fact, most countries publish price of some of the basic items on an annual basis. For items not covered by such publications, suitable price deflators from the producer price index (PPI) can be considered. Changes in wages for construction labor are also frequently available from national sources which can be used in the construction of PPPs.
Machinery and equipment (M&E) is a more complex aggregate to deal with. Currently, PPPs are compiled using prices collected for a global list of products included in the M&E list. However, difficulties arise in the practical implementation of this approach. For purposes of annual compilation, reliance could be made on the fact that PPPs for M&E are usually to the exchange rate as most of the items are traded and frequently imported. Blades (2013) examined this possibility and it offers a way of compiling adjustment factors to bring exchange rates in line with PPPs. This is an option that should be seriously explored.

4.5 Global Core List for Regional Linking

The global core list for household consumption has significant overlap with product lists used in different regions. So a blend of rolling price-survey and annual price-survey approach can be used in improving the quality of the linking factors. As the global core product lists and linking is done at the Global Office at the World Bank, it may be feasible to make use of alternative sources, such as internet and scanner data sources, of prices for products in the global core list.

4.6 Asymmetric data availability

Until the process of annual compilation of PPPs is set in place, it is quite possible that data available for PPP computation differs significantly and qualitatively across different countries in different regions. The current situation is that the level of disaggregation of CPI and other national price deflators differ. In the extreme case, only the GDP deflator may be available for a country where as deflators are available at a much higher levels of disaggregation.

In the case where availability and quality of data vary across countries, our approach is to use all the available information rather than to operate at the level of the lowest common denominator.

Section 5. Approach for updating PPPs

The starting point in our updating approach is the data for all the regions provided by the ICP Global Office and Eurostat:

1. Basic heading (BH) expenditure levels for the years 2011, 2012 and 2013
2. BH PPPs (linked across regions) for 2011 from ICP 2011

3. BH PPPs for 2012 and 2013 for Western Asia based on extrapolation using detailed inflation information at ESCWA
4. BH PPPs for 2012 and 2013 for Eurostat countries from its rolling benchmark program.
5. National Accounts (NA) deflators for GDP and main expenditure aggregates and Consumer Price Index (CPI) figures for overall household consumption and by COICOP category for 2011, 2012, and 2013.
6. Exchange rate and population information for 2011, 2012 and 2013.

5.1 Data validation

The framework in Section 3 made clear under what circumstances we may expect a divergence between relative inflation rates and changes in PPPs driven by differences in expenditure shares and relative price movements. Those circumstances assumed that the same index number formula is used for inflation calculation and PPP computation and that the expenditure shares match. In practice, these assumptions do not hold. Budget shares in the computation of inflation are typically not updated annually, they may differ systematically from National Accounts expenditure shares, for instance due to differences in the treatment of housing or of financial intermediation services indirectly measured (FISIM). Furthermore, ICP PPPs are computed using the GEKS index number approach, based on bilateral Fisher indexes, while inflation and National Accounts deflators are more typically computed based on fixed or chained-Laspeyres indexes. Moreover, given the varying degree of statistical capacity amongst the 177 ICP countries, it is conceivable that inconsistencies occur in some series for some countries.

It is beyond the scope of this report to comprehensively account for every country's data and methods, but we can do a comparison between observed overall inflation rate and a counterfactual inflation rate computed using detailed inflation rates and NA expenditures. There are two datasets with detailed inflation information, namely the NA deflators for major expenditure categories (household consumption, gross capital formation, etc.) and the CPI inflation rates by COICOP category, so we run two comparisons. In the first comparison, we compute expenditure shares within household consumption by COICOP category[8] and use a Törnqvist index number to compute overall inflation rate:

$$\ln P_{t,t+1} = \frac{1}{2}\sum_{i=1}^{N}(s_{it} + s_{it+1})\left(\ln p_{i2}^{t+1} - \ln p_{i2}^{t}\right) \tag{16}$$

Comparing the counterfactual inflation rates based on equation (16) to the official overall inflation rates shows differences in every country, as expected given the preceding discussion. Table 1 presents summary statistics for the difference between actual reported inflation and the counterfactual based on equation (16). The median difference is zero, indicating that there is no systematic bias in the counterfactual measure. The distribution is wide however. The 75th percentile number for 2012 indicates that for a quarter of the countries, we find a counterfactual inflation rate that is 0.20 percentage points faster than the observed inflation rate and for 10 percent of the countries, the differences even exceed 0.55 percentage point. At the extremes of the distribution, there are countries for which the inflation rate differs by a full percentage point or more. Given median actual observed inflation of 3.5 percent in 2012 and 2.5 percent in 2013, the difference can be large.

Table 1: Difference between actual and counterfactual CPI inflation in 2012 and 2013 (in %)

	2012	2013
10TH PERCENTILE	-0.75	-0.61
25TH PERCENTILE	-0.37	-0.25

---

[8] Data coverage of COICOP categories varies, but we compute expenditure shares ignoring COICOP categories without inflation rate coverage. We include all 135 countries with data on the overall CPI and 7 or more COICOP-level inflation rates.

MEDIAN	0.00	-0.00
MEAN	-0.12	0.05
75ᵀᴴ PERCENTILE	0.20	0.23
90ᵀᴴ PERCENTILE	0.55	0.61

Notes: counterfactual inflation is computed based on equation (16), negative numbers means that counterfactual inflation is lower than actual reported inflation.

We performed a similar exercise for the NA deflators, comparing the actual reported GDP deflator change to the deflator change computed based on the underlying expenditure categories. Since the balance of exports and imports is frequently a negative number, we compute our counterfactual index as a chained-Laspeyres rather than a Törnqvist index. Table 2 reports the results of this comparison. The differences are notably larger than for CPI inflation, with even the median country showing a difference of 0.38 and 0.08 percentage point. These larger differences are surprising, since 1) NA expenditure shares are used in the computation of the GDP deflator as well and 2) many countries rely on a chained-Laspeyres index to compute their GDP deflator.

Table 2 Difference between actual and counterfactual GDP deflator change in 2012 and 2013 (in %)

	2012	**2013**
10ᵀᴴ PERCENTILE	-3.33	-1.89
25ᵀᴴ PERCENTILE	-0.71	-0.61
MEDIAN	0.38	0.08
MEAN	0.28	0.47
75ᵀᴴ PERCENTILE	1.30	1.34
90ᵀᴴ PERCENTILE	2.33	3.01

Notes: see Table 1.

Again, a full accounting of the differences is beyond the scope of this report. A more detailed comparison for the Netherlands showed that the deflator for the balance of exports and imports in the provided dataset differs from the deflator that can be derived from Statistics Netherlands data. Similarly, a comparison of CPI data released by the Indonesian statistical bureau (BPS) suggests that the source used in the provided data (Haver Analytics) may not be fully reliable.

More generally, the comparisons in Table 1 and 2 suggest this is a useful data validation framework to assess the reliability of the input data. Ideally, this can help point to and identify errors or other shortcomings or to decide amongst alternative sources for the same country. For our subsequent analysis, we will rely on the data 'as is', but in interpreting the results, it is useful to bear in mind that there are not just conceptual reasons for observed differences between relative inflation rates and changes in PPPs but also data-related reasons.

5.2 Data preparation

Given the scope of the available data, it is feasible to cover 172 countries. These include all countries that participated in the global GDP-level comparison of ICP 2011, except for those countries for which no inflation data were available[9] and Cuba, due to suppressed expenditure information.[10] The data allow us to distinguish 151 basic headings, which is less than the 155 in ICP 2011 because the constituent elements of the balancing items were not separately available.[11] Since separate PPPs for those constituent elements were not available in ICP 2011, this does not lead to a loss of precision.

---

[9] Excluded for this reason are Bonaire, Curacao, Montserat, Sint Maarten, Turks and Caicos Islands and the British Virgin Islands.

[10] The dual-participation countries, Egypt, Russia, and Sudan, are counted only once in the list of 172.

[11] The combined balancing items are 1) balance of expenditures of residents abroad and expenditures of non-residents in the economic territory, 2) changes of inventories, 3) acquisitions less disposals of valuables, and 4) balance of exports and imports.

The data for Western Asia (item 3) and Eurostat countries (item 4) were matched to those 151 basic headings and, in the case of Eurostat, this involved aggregating over more detailed expenditure categories.

Expenditure data at the basic heading level has been estimated in a nearly comprehensive fashion. The first exception is the division of government spending on health and education into labor compensation and other costs. We use the distribution of expenditure across cost categories from 2011 to estimate spending in 2012 and 2013. The second exception is expenditure data for Aruba, Anguilla and (in 2013) the Cayman Islands. For those cases, we use the 2011 expenditure levels for 2012 and 2013.

COICOP-level CPI data is available for many countries, but coverage of the consumption categories is incomplete, ranging from fairly complete for food and non-alcoholic beverages to more limited for health or education. To ensure the 151 basic headings for all 172 countries can be extrapolated from 2011 to 2012 and 2013, we used the following approach. First, for household consumption, the COICOP-level CPI data were used if available; if not, the overall CPI was used; if not, the household consumption deflator was used. For the other expenditure categories (government consumption and gross capital formation), we used national accounts deflators if available; if not the GDP deflator; if not the overall CPI was used. For the BH where the exchange rate is used in ICP, such as the balance of exports and imports, we also use the exchange rate in 2012 and 2013.

5.3 Compilation of 2012 and 2013 global PPPs
At this stage, we have a dataset with expenditure levels for 151 basic headings and 172 countries for 2011, 2012 and 2013; globally-linked PPPs for all basic headings in 2011 and regional PPPs for Eurostat countries and Western Asian countries in 2012 and 2013; and estimates of inflation rates for all basic headings.
The first step in the compilation of global PPPs for ICP aggregates is to extrapolate the linked 2011 BH PPPs using the rate of inflation of each BH in each country relative to the United States. Note, first, that the differing availability of inflation data can mean that a BH PPP for one country is extrapolated using the GDP deflator (relative to the more detailed inflation rate in the US), while for another country more detailed data is used. Also note that the extrapolation of linked BH PPPs is equivalent to: i) extrapolating the within-region BH PPPs using the inflation rate of each country relative to the regional base country; ii) extrapolating the item prices of the global core list (GCL) using the inflation rate of the corresponding BH; and iii) deriving new BH linking factors based on weighted-CPD models. This equivalence is due to the fact that all item prices within a BH are extrapolated with the same inflation rate.

The second step is to replace the extrapolated linked BH PPPs for Eurostat and Western Asian countries by the actual PPPs for 2012 and 2013. These actual PPPs are integrated by multiplying the Western Asian PPPs by the Oman PPPs (i.e. the Western Asia linking factors) and the Eurostat PPPs by the Germany PPPs. The next steps follow the ICP 2011 Global Report (World Bank, 2015, pages 211 and 212) in linking across regions at aggregate levels. So for the Eurostat/OECD, Africa, Asia-Pacific, Latin America and Western Asia regions, we apply the CAR method. Then we enforce fixity of Eurostat PPPs within the Eurostat/OECD region, link the CIS countries through Russia, link the Caribbean region to the global comparison through a CAR link with Latin America, link Georgia through Armenia, and Iran through Turkey. The final step is to get a single PPP estimate for the dual-participation economies (Egypt and Sudan) as the geometric average of the global PPPs from both regions.[12] The end result is a set of global PPPs for all ICP aggregates for 2012 and 2013.

To place the estimates for 2012 and 2013 in context, we also compute PPPs for all ICP aggregates using so-called 'global extrapolation', i.e. extrapolating the PPPs at an aggregate level using aggregate inflation figures (bearing in mind the results of the data validation exercise discussed

---

[12] Russia is also a dual-participation economy, but since CIS is linked to Russia directly, no averaging is needed.
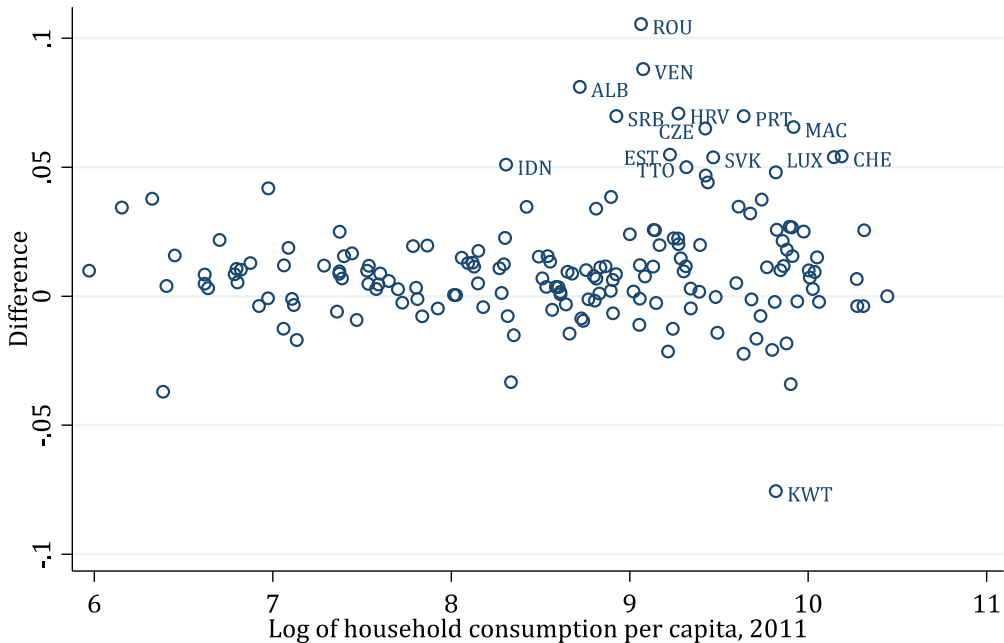
above). Global extrapolation is the current approach in the World Development Indicators, but this approach cannot adequately account for changing expenditure patterns and PPPs at lower levels of aggregation.

Section 6: Results – Extrapolated PPPs for 2012 and 2013

Figure 1 compares the household final consumption expenditure PPPs in 2013 based on the detailed updating approach discussed in the previous section with the PPPs based on the global extrapolation approach, using the overall CPI for updating. With the ICP 2011 results as the starting point for both PPPs, the figure illustrates the differences that accumulate in two years of using global updating instead of using detailed inflation figures and changing expenditure figure.

The first observation from this figure is that the average difference is close to zero and that differences are not systematically related to (log) consumption per capita. Note that the United States is taken as the base country for the PPPs and that the difference measure we define is not base-country independent. However, while this affects the average difference level, the relationship with (log) consumption per capita is unaffected by the base country choice.

Figure 1, Difference between global extrapolation and detailed updating of PPPs: household final consumption expenditure, 2013.

Notes: Global extrapolation uses the change in CPI (or household consumption deflator) relative to the US to estimate PPPs for subsequent years. Detailed extrapolation uses the CPIs at the most detailed level to extrapolate basic-heading level PPPs, reflects changing basic heading expenditure levels, and links regions above basic heading level according to the ICP 2011 methodology. This also includes benchmark PPPs from Eurostat and the more detailed extrapolation for Western Asia.
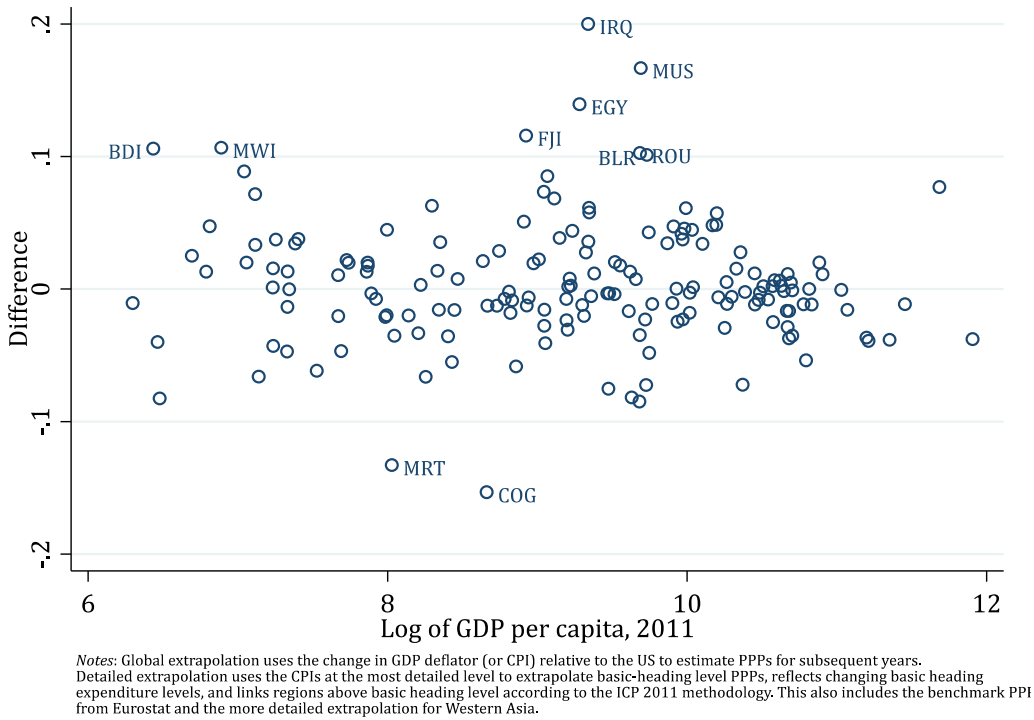
The second observation is that for the great majority of countries, the difference between the two methods is less than 5 percent. While not drastic, the absolute difference for the median country is 1 percent, or 0.5 percent per annum. This implies that a 3 percent difference can easily open up if ICP rounds are 6 years apart, as between ICP 2005 and ICP 2011. The results from this comparison are also a useful diagnostic tool: an earlier set of results showed a large difference for the Comoros, but these were traced to a 63 percent annual price decline for one consumption category and a 123 percent increase for another. Replacing these price trends by the trend in the overall CPI led to a much smaller difference.

The third observation is that nearly all countries for which differences are larger than 5 percent (in absolute sense) are in the Eurostat and Western Asia regions. These are the only two regions with PPP estimates for 2012 and 2013, from Eurostat's rolling-benchmark approach and from Western Asia's

13

more detailed updating efforts. For comparison, Appendix Figure 1 shows the results when ignoring the new benchmark figures for those two regions and most countries that were above the 5 percent line (indicating that global extrapolation leads to a faster increase in the PPP than the more detailed updating) show smaller differences. This result is consistent with the earlier finding of Inklaar and Timmer (2013a) for Eurostat countries that new survey results can lead to notable changes, perhaps reflecting sampling error in the surveys. The case of Kuwait seems to reflect some degree of re-benchmarking in Western Asia, as the PPPs for water supply (basic heading code 1104411), miscellaneous dwelling services (1104421) and electricity (1104511) change much more (130–570%) than seems warranted by relative inflation.

Figure 2 shows the same comparison global extrapolation and detailed updating at the GDP level. Note that the scale differs between the two figures and that the large majority of countries show differences less than 10 percent, rather than less than 5 percent as was the case for household final consumption expenditure. One obvious candidate for explaining these larger differences is the trade balance, since the exchange rate is used to convert this to a common currency. However, Appendix Figure 2 shows that differences for domestic absorption (consumption plus investment) are very similar to the GDP comparison in Figure 1.

Figure 2, Difference between global extrapolation and detailed updating of PPPs: GDP, 2013.

Notes: Global extrapolation uses the change in GDP deflator (or CPI) relative to the US to estimate PPPs for subsequent years. Detailed extrapolation uses the CPIs at the most detailed level to extrapolate basic-heading level PPPs, reflects changing basic heading expenditure levels, and links regions above basic heading level according to the ICP 2011 methodology. This also includes the benchmark PPI from Eurostat and the more detailed extrapolation for Western Asia.

This is a useful point to recall the data validation exercise from the previous section. That exercise showed that the GDP deflator was often quite different from a chained-Laspeyres aggregate of NA deflators, pointing to inconsistencies between the two. Furthermore, since those inconsistencies were larger for the GDP deflator than for the analogous comparison of CPI inflation, this could account for the larger differences in Figure 2 compared to Figure 1. Given the experimental scope of this updating exercise, we note this as one possible explanation for the observed differences.

Section 6. Annual Compilation of PPPs

From the earlier work on inconsistency between benchmarks (Dalgaard, 2002; McCarthy, 2013; Inklaar and Timmer, 2013; Deaton and Aten, 2016; and Inklaar and Rao, 2016) it is desirable to

compile PPPs on an annual basis. The analytical framework discussed in Section 2 shows that where extrapolations are needed, it is best if the extrapolations are done at the basic heading level. In sections 5 and 6, the paper demonstrated the feasibility of compiling PPPs for 2012 and 2013 making use of primary data collected for 2011 and price deflator data available from the participating countries.

In this section we provide a blueprint for the compilation and publication of PPPs and real GDP aggregates on an annual basis. The main objective here is to identify the major steps involved. Once the steps are agreed upon, administrative arrangements and the involvement of regions in this process can be identified.

We our focus in this section is the compilation of PPPs for the years 2017 to 2020. However, the year 2017 is given special treatment as there has been a six-year gap from the last benchmark in 2011 and the process is then extended to cover the ensuing years 2018 to 2020.

6.1 Rolling survey approach for benchmarks versus annual compilation of PPPs

We make a clear distinction between the rolling survey approach to benchmark comparisons and annual compilation of PPPs. In the three-year rolling price survey approach, after completion of a benchmark comparison in period t, price data necessary the next benchmark in year $t+3$ are collected using a series of rolling surveys. Under this approach, one-third of the items are priced in each of the three years so that all items are have fresh survey prices collected at least once during the three year period. The prices collected in years $t+1$ and $t+2$ are updated to the benchmark year $t+3$ using appropriate price deflators. All the data collected through these three years are used compiling the benchmark comparisons in $t+3$. Details of the rolling benchmark approach can be found in Eurostat (2012).

The annual compilation of PPPs is a different exercise. First, each year in annual compilation is a like a benchmark comparison. There is no special status attached to any single year. Second, the survey and methodological approaches used for the compilation in each year are essentially the same except for the introduction of innovations in methodology for international comparisons. There is a continued focus on the refinement of methodology for comparison-resistant sectors like education and health sectors and also problem areas like dwellings and construction. Finally, as annual comparisons are greatly affected by the availability of information on national accounts which are themselves subject to revision it may be necessary to have a well-defined revisions policy on PPPs and real GDP comparisons compiled annually.

6.2 Framework for Annual Compilation of PPPs

The current proposal for annual compilation represents a major departure from the current practice of undertaking international comparisons of prices and real expenditures once in five to six years and then to provide extrapolations to non-benchmark years.[13] Annual compilation of PPPs eliminates the need for interpolation and extrapolation as it has been the past practice. The following are the main components in our framework for the annual compilation of PPPs.

- Benchmark comparison for 2017

- Compilation of annual series from 2018 onwards

- Annual series for the years 2012 to 2016 – between the 2011 ICP benchmark and the 2017 benchmark comparison recommended by the UN Statistical Commission

- Revisions and publication policy for annual price and real expenditure comparisons.

---

[13] Extrapolation to non-benchmark years is an exercise that has attracted considerable interest. Apart from the World Bank approach for use in the *World Development Indicators*, the PWT (Feenstra et al, 2015) uses as interpolation strategy whereas the UQICD uses an econometric approach (Rao, Rambaldi and Doran, 2010, 2015) for extrapolations.

6.3 Benchmark comparison for 2017

The benchmark comparison for 2017 assumes special significance due to the recommendation of the UN Statistical Commission (UNSC) to move the ICP towards an annual program instead of a five-year benchmark program. Thus the 2017 ICP benchmark cannot be implemented like a standard ICP benchmark where all the price and national accounts data are collected for the benchmark year. A further complicating factor is that the preparation period for this exercise is short as it is already well into 2016. Taking these factors into account we propose the following approach that is a mixture of the standard ICP approach and the rolling survey approach.

Table 3: Survey approach for the 2017 ICP Benchmark

Aggregates	2017	2018
Household Final Consumption		
Food and nonalcoholic beverages		⇐ X
Alcoholic beverages, tobacco and narcotics		⇐ X
Clothing and Footwear		⇐ X
Dwellings	X	⇐ X
Water, electricity, gas and other fuels	X	
Furnishings, household equipment and routine household maintenance		⇐ X
Health	X	
Transport	X	
Communication	X	
Recreation and culture	X	
Education	X	
Restaurants and Hotels		⇐ X
Miscellaneous goods and services		⇐ X
Government Final consumption		
General consumption		⇐ X
Compensation of employees	X	
Gross Fixed Capital formation		
Machinery and equipment	X	
Construction	X	
Exports	X	
Imports	X	
Global Core List	X	
Household Consumption		⇐ X[*]
Other – Government, GFCF	X	
National Accounts Expenditure Share data	**MA_15,16,17**	**MA_16,17,18**
National Accounts Expenditure Data	**X**	**X**

(*) indicates that the schedule for GCL for Household Consumption follows the same as the schedule for
different sub-aggregates for household consumption.

In Table 3, the symbol "**X**" means that price surveys need to be conducted in that year. The symbol "⇐ **X** " indicates that price data are collected in 2018 and then extrapolated backwards to provide prices for compiling PPPs in 2017.

The survey strategy outlined in Table 1 is based on the following considerations:

16

- It is not possible to conduct a full-scale ICP exercise in 2017. Previous experience suggests that conducting an ICP round requires considerable lead time and time for processing and finalizing the results. For example, work on the 2011 ICP began in 2009 and concluded early in 2015.
- Given the timing of the UNSC's recommendations and the fact that work on the 2017 comparisons has just begun, it is not feasible to conduct any price surveys in 2016 which could be extrapolated to 2017.
- The survey framework in Table 3 is designed to make it feasible to collect data required for the 2017 benchmark comparisons over 2017 and 2018.
- Given the long gap since the last benchmark, the proposal here allows for a recalibration of the list of goods and services to be priced in household consumption and also items included in the Global Core List.
    o As CPI movements track of food, beverages and tobacco; and clothing footwear, price collection for these aggregates are scheduled for 2018 and the resulting prices extrapolated backwards to 2017 using movements in CPI for these item groups.
    o Dwellings is an item that requires improvements in methodology and given the time gap, it is hoped that regions like Asia-Pacific and Africa are in a position to provide better quantity indicator data to replace the current approach of "reference quantity approach" by a more direct measure. Consistent with the principle of not introducing major methodological innovations from year to year, it is recommended that the methodology for dwellings is refined and implemented in 2017 and 2018. We have symbol "**X**" in both 2017 and 2018 to mean that work on collection of indicator data must start in 2017 and possibly spill over to 2018.
    o The water, electricity, gas and other fuels data are usually easy to obtain and are usually available from CPI price data collection and therefore flagged to be collected in 2017.
    o Furnishings; restaurants and hotels; and miscellaneous goods and services are marked for survey in 2018.
    o The aggregates of transport; communication; recreation and culture are set for price collection in 2017. As there are rapid changes in goods and services belonging to these aggregates, it is important that this phenomenon is captured through calibration of products lists on an annual basis and price data are collected accordingly.
    o Health and Education are set for price collection in 2017 as there may have been major shifts in goods and services since the last price collection.
- Under government expenditure, general consumption is included in 2018 and government compensation and the related productivity adjustments necessary are included in 2017.
- Machinery and Equipment is included in 2017 as it is a component of GCL items. Modifications to the product list based on the experience of the participating regions need to be introduced before the 2017 price surveys. There is scope for M&E price surveys to be conducted over the two-year period 2017 and 2018.
- The GCL items are classified into household consumption items and the remaining items are classified as "other". The proposal, consistent with the treatment of household consumption, is to collect prices for household consumption aggregates in 2018 and extrapolate backwards. However, the survey framework here needs to be synchronized with the framework used for sub-aggregates in household consumption.
- Exchange rates are used for exports and imports and therefore collected each year.
- National accounts data are required for two purposes.
    o First is the expenditure share data required for aggregation. We propose that the aggregation in 2017 be based on the moving average of expenditure shares in 2015, 2016 and 2017. This is designed to smooth any errors present in expenditure share data.
    o The second purpose is to compile real expenditure data and per capita volume measures at the GDP level and for sub-aggregates. For this purpose, it is necessary to make use of national accounts data for 2017. There could be some delay in getting this data and hence preliminary accounts for 2017 may be available only some time in 2018. The timing of availability of national accounts data and the revisions policies adopted in different countries will dictate the revisions policy for PPPs and real expenditures from ICP.

17

6.4 Compilation of annual series from 2018 onwards

The following table show the general survey framework for the compilation of annual PPPs.

Table 4: Survey approach for Annual Compilation of PPPs, 2018 and beyond

Aggregates	2018	2019	2020
Household Final Consumption			
Food and nonalcoholic beverages	S $\Rightarrow$	E_2018	E_2018
Alcoholic beverages, tobacco and narcotics	S $\Rightarrow$	E_2018	E_2018
Clothing and Footwear	S $\Rightarrow$	E_2018	E_2018
Dwellings	E_2017	S	E_2019
Water, electricity, gas and other fuels	E_2017	S	E_2019
Furnishings, household equipment and routine household maintenance	S	E_2018	E-2018
Health	E_2017	S	E_2019
Transport	E_2017	S	S
Communication	S	S	S
Recreation and culture	E_2017	E_2017	S
Education	E_2017	S	E_2019
Restaurants and Hotels	S	E_2018	$\Leftarrow$ X
Miscellaneous goods and services	S	E_2018	$\Leftarrow$ X
Government Final consumption			
General consumption	S	E_2018	$\Leftarrow$ X
Compensation of employees	E_2017	S	E_2018
Gross Fixed Capital formation			
Machinery and equipment	E_2017	E_2017	S
Construction	S	E_2018	
Exports	S	S	S
Imports	S	S	S
Global Core List			
Household Consumption	S	E-2018	E_2018
Other – Government, GFCF	E_2017	E_2017	S
National Accounts Expenditure Share data	MA_16,17,18	MA_17,18,19	MA_18,19,20
National Accounts Expenditure Data	S	S	S

Table 4 shows a feasible plan for the compilation of PPPs for the years 2018 to 2020 on an annual basis. The symbol "S" means that there will be a price survey for that particular aggregate where as "E" stands for extrapolation from a designated year. According to this plan, extrapolations are done for a maximum of two years. Aggregates such as "communication" and "transport" are scheduled for more frequent surveys as new products are introduced constantly. National accounts expenditure share data used in aggregation are based on a three-year moving average which will smooth the share data. However, the actual national accounts expenditure aggregate data are collected for each year separately. Timely availability of national accounts data and frequent revisions by the national statistical offices will require careful handling of these data and a proper "revisions" policy for annual PPPs and real expenditure aggregates. *The schedule given in Table 4 is indicative and needs to be finalized after extensive consultations with all the stakeholders.* We also note a feature of the proposal here whereby only forward extrapolation of price data is used. This is in contrast to the general

rolling-price survey approach implemented over a three-year period. The forward extrapolation ensures timely release of ICP results on an annual basis.

6.5 Dissemination of results – Preliminary, Revised and Final Results

The following is an illustrative guideline to the release of results for the 2017 benchmark and for the subsequent years.

Table 5: Timeline for release of ICP results

Benchmark Year	Preliminary Results	Revised Results	Final Publication
2017 Benchmark Year	July, 2019	December, 2019	July, 2020
2018 Benchmark Year	July, 2019	December, 2019	July, 2020
2019 Benchmark Year	July, 2020	December, 2020	July, 2021
2020 Benchmark Year	July, 2021	December, 2021	July, 2022

The release and finalization of results for the 2017 benchmark year will take a little longer as the results will be based on surveys conducted in 2018. Preliminary release of results by July 2019 will require a streamlined approach to the computation and validation of results. Once the process of annual compilation is institutionalized, then it would be possible to publish final results for each years within 18 months after the completion of collection of primary data for aggregation. The timetable set here may be ambitious but it is a goal that needs to be achieved if the results are to be released in time for use by the World Bank, IMF and other organizations and researchers. *In order to implement this timetable, it is necessary to articulate and establish an agreed policy on revisions and finalization of results.*

- PPPs and real expenditure released as a part of *preliminary results* are both subject to revision – revisions could be due to small changes to price data or to the national accounts weights used in aggregation.
- At the stage of release of *revised results*, the PPPs published at this stage have to be final and not subject to change as there is no expectation of any change in prices or weights. However, real expenditure data could be subject to change due to revisions undertaken by the national statistical offices.
- The *final results* published are not subject to further change unless there are exceptional circumstances whereby some participating countries may have changed their national accounts data.

6.6 Annual Series for years between the 2011 and 2017 benchmarks

As the results from the 2017 benchmark are not expected until the middle of 2019, it is essential to make arrangements for the release of extrapolated results for the years between the two benchmarks.

- Extrapolations for the years 2012 to 2014: We recommend that extrapolation of results for 2012 to 2014 is undertaken using the methodology described in this paper. The basic principle is to make use of all the possible information available from all sources in any given year and, where necessary, use price data extrapolated from 2011 constructed using detailed CPI and national accounts deflators. This approach makes it possible to accommodate price survey data available in some regions which have been undertaking extrapolation activities. *Essentially extrapolations to 2012 to 2014 use ICP 2011 price data; CPI and National Accounts Deflators; annual data on national accounts; any additional updated price survey data from the regions.*
- We recommend a slightly different approach for the years 2015 and 2016. As these years are further apart from 2011, we believe that a backward extrapolation from 2017 benchmark is more appropriate for these two years. The backward extrapolated data from 2017 combined with any survey data available from the regions would be combined to yield PPPs and real expenditures for 2015 and 2016.

- Given that revised results for 2017 are not available until December, 2019, results for 2015 and 2016 can be released only in 2020. There may be advantages in simply extrapolating from 2011 and making use of any intermediate price survey data that may be available from the participating regions. This means that results for 2015 and 2016 can be made available by mid to late December, 2017. This is an executive decision to be made by the implementing agency at the World Bank. Using this approach means that it would be possible to provide PPPs for publication in the *World Development Indicators* in a timely manner. In the absence of results for 2015 and 2016, it is possible that users may simply extrapolate results at the GDP level which is a practice that needs to be avoided. A compromise would be to base initial estimates of 2015 and 2016 on extrapolations from 2011 and revise these when 2017 benchmark results are available.

Section 7. Conclusions

The main objectives of the paper are to: (i) examine the problem of extrapolation of PPPs from the 2011 benchmark to provided extrapolated data for the years 2012 to 2016; (ii) provide a blueprint for the compilation of PPPs for the 2017 benchmark years; and (iii) to provide a framework for compilation of annual PPPs. The paper provides an analytical framework that underpins the extrapolation approach. The main conclusion is that in order to minimize the discrepancies between benchmark comparisons and temporal movements in prices, extrapolations must be undertaken at the lowest aggregate possible where it is likely that movements of prices of products within the group are quite similar in magnitude over time. The paper adheres to the principle of making use of all the data available at any given point of time to construct PPPs instead of anchoring the price and real expenditure comparisons on the country or region with least amount of available data. We demonstrate the feasibility of this approach and construct PPP extrapolations for the years 2012 and 2013 making use of price comparison data available from 2011 and combining these with additional price data available from regions and with the CPI and National Accounts deflators available at the most disaggregated level. The last section of the paper outlines a feasible plan for the implementation of the 2017 benchmark comparisons and for the subsequent compilation of PPPs for the years 2018 to 2020. The paper also discusses the dissemination plan which releases preliminary results which are then revised and finalized within 18 months from the publication of the preliminary results. We believe that the proposals presented here provide a foundation to the methodology for the annual compilation of PPPs at the World Bank.

References

Asian Development Bank (2013), *Purchasing Power Parities and Real Expenditures*, Manila.

Balk, B.M. (2008), *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference.* New York: Cambridge University Press.

Dalgaard, E. and H.J. Sorensen (2002), "Consistency between PPP benchmarks and national price and volume indexes", OECD Meeting of national accounts experts, Working Paper.

Deaton, Angus (2012), "Consumer price indexes, purchasing power parity exchange rates, and updating", Mimeographed, Princeton University, Princeton, NJ.

Deaton, Angus and Alan Heston, 2010. "Understanding PPPs and PPP-Based National Accounts," American Economic Journal: Macroeconomics, American Economic Association, vol. 2(4), pages 1-35, October.

Deaton, Angus and Bettina Aten, 2016. "Trying to understand the PPPs in ICP 2011: why are the results so different?" forthcoming *American Economic Journal: Macroeconomics*.

ECE-ILO (2010), Consumer Price Index Manual: Theory and Practice, Geneva.

Feenstra, R.C., Hong Ma, J. Peter Neary and D.S. Prasada Rao, 2013. "Who Shrunk China? Puzzles in the Measurement of Real GDP," Economic Journal, Royal Economic Society, vol. 123(12), pages 1100-1129, December.

Feenstra, R.C., R. Inklaar and M.P. Timmer (2015), "The Next Generation of the Penn World Table", American Economic Review, 105(10): 3150-82.

Inklaar, R. and M. P. Timmer (2013a), "A Note on Extrapolating PPPs**",** Annex D to Chapter 18 in World Bank edited book *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program – ICP*, Washington DC.
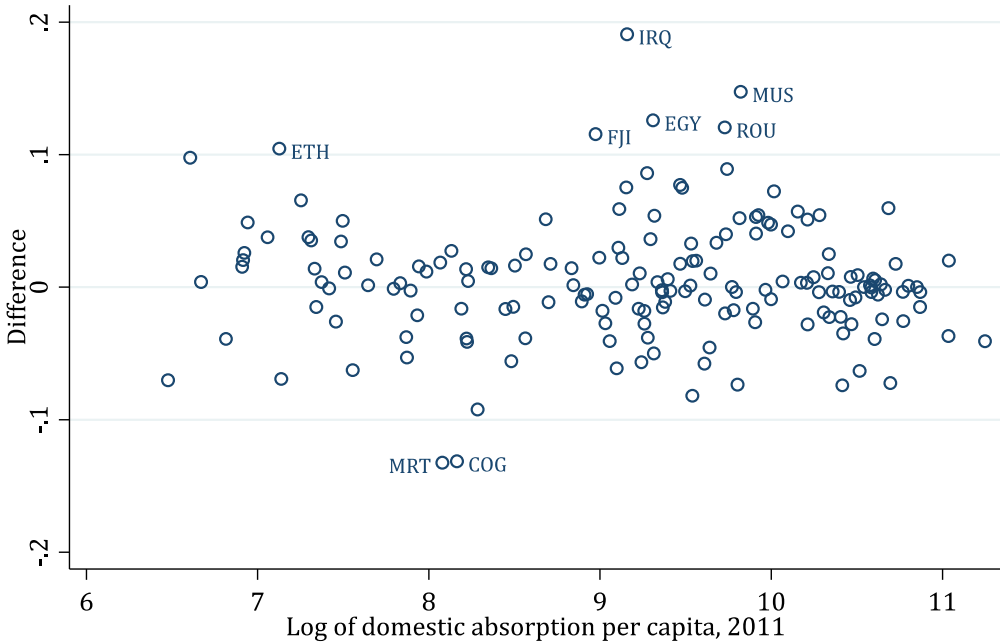
Inklaar, R. and M.P. Timmer (2013b), "Productivity Adjustment for /government Services PPPs: Alternatives and Proposal for ICP 2011", Groningen Growth and Development Centre, University of Groningen, Groningen.

Inklaar, R. and D.S. Prasada Rao (2016), "Cross-country income levels over time: did the developing world suddenly become much richer?" forthcoming in *American Economic Journal: Macroeconomics*.

McCarthy, P.J. (2013a), "Extrapolating PPPs and Comparing ICP Benchmark Results", Chapter 18 in World Bank edited book *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program – ICP*, Washington DC.

McCarthy, P.J. (2013b), "Governance Structure of ICP 2005", Chapter 2 in World Bank edited book *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program – ICP*, Washington DC.

Rao, D.S. Prasada (2013), "The Framework of the International Comparison Program", Chapter 1 in World Bank edited book *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program – ICP*, Washington DC.

Rao, D. S. P., Rambaldi, A.N. and Doran, H.E. (2010), "Extrapolation of purchasing power parities using mul- tiple benchmarks and auxiliary information: A new approach." Review of Income and Wealth, 56 Suppl 1: S59-S98.

Rao, D.S.P, A. N. Rambaldi, K.R. Ganegodage, L. T. Huynh, H. E. Doran, (2015) " UQICD v2 User Guide " School of Economics Discussion Paper No. 534 October 2014, School of Economics, The University of Queensland. Australia.

World Bank (2013), *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program – ICP*, Washington DC.

World Bank (2015), *Purchasing Power Parities and the Real Size of World Economies – A Comprehensive Report of the 2011 International Comparison Program*, Washington DC.

21

Appendix Figure 1, Difference between global extrapolation and detailed updating of PPPs: household final consumption expenditure, 2013; no new benchmark information.
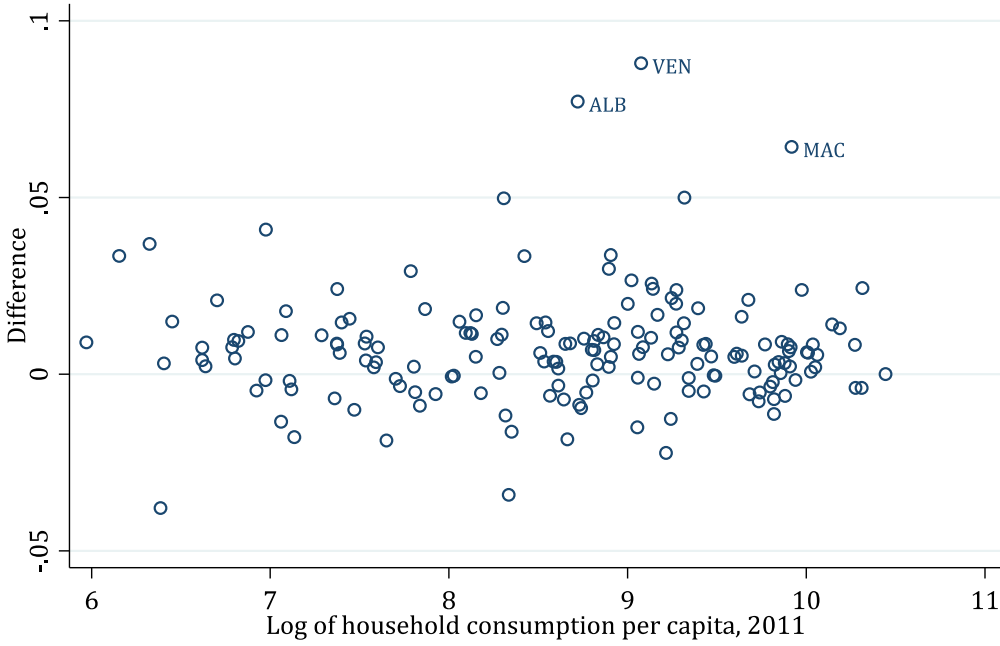
Notes: Global extrapolation uses the change in CPI (or household consumption deflator) relative to the US to estimate PPPs for subsequent years. Detailed extrapolation uses the CPIs at the most detailed level to extrapolate basic-heading level PPPs, reflects changing basic heading expenditure levels, and links regions above basic heading level according to the ICP 2011 methodology. This also includes the benchmark PPI from Eurostat and the more detailed extrapolation for Western Asia.

Appendix Figure 2, Difference between global extrapolation and detailed updating of PPPs: domestic absorption, 2013.

Notes: Global extrapolation uses the change in GDP deflator (or CPI) relative to the US to estimate PPPs for subsequent years. Detailed extrapolation uses the CPIs at the most detailed level to extrapolate basic-heading level PPPs, reflects changing basic heading expenditure levels, and links regions above basic heading level according to the ICP 2011 methodology. This also includes the benchmark PPI from Eurostat and the more detailed extrapolation for Western Asia.