

# International Comparison Program

---

## Understanding Changes in Methodology between the 2005 and 2011 International Comparison Programs

Paul McCarty, TAG Chair and  
Fred Vogel, TAG Co-Chair



**10<sup>th</sup> Executive Board Meeting**  
January 24, 2014  
Washington, DC

# Understanding Changes in Methodology between the 2005 and 2011 International Comparison Programs

## Executive summary

This paper provides an overview of the technical decisions that underlie the results of the 2011 International Comparison Program (ICP) and concludes that there are no further steps or computations that would have any substantive impact on the final results. It also provides a review of methodological changes that can affect the comparison of the 2011 results with the 2005 results. The 2005 ICP produced estimates of purchasing power parities (PPPs), real expenditures and price level indexes for 146 countries around the world, making it the largest statistical exercise undertaken on a global basis. A number of new techniques were introduced in the 2005 ICP, and the detailed data available from that round enabled an in-depth analysis to be undertaken of not only the results but also the impact of various methods that were employed, particularly those used to link the PPPs for the six “regions”<sup>1</sup> to provide consistent estimates for all 146 countries.

The outcomes of the comprehensive analysis of the methods used in the 2005 ICP were described in detail in the “2005 ICP book” – “Measuring the real size of the world economy” by the World Bank (2013)<sup>2</sup>. A number of areas were highlighted in this book as being candidates for improved methods – both statistical techniques and data collection – and the Technical Advisory Group (TAG) recommended that a number of them should be implemented in the 2011 ICP.

The benefits of implementing new methods to overcome shortcomings in previously-used procedures are fairly clear, particularly as regards ensuring that the 2011 inter-country comparisons are as sound as possible. However, the downside of improving the methods is that it introduces inconsistencies between the successive sets of benchmark results for 2005 and 2011. Chapter 18 in the 2005 ICP book describes in some detail the reasons why changes between the 2005 and 2011 ICP results would not be consistent with the changes in the time series national accounts of the countries involved, even if all the data and methods were consistent between the two benchmark years. In practice, though, the methods used in the 2011 round have been improved in a number of ways as a result of the detailed analysis of the 2005 results, which adds an additional element of complexity when trying to explain changes between the 2005 and 2011 benchmarks.

The most significant methodological change made for ICP 2011 was how countries are linked across regions. This change was complex because so many different factors were involved. Instead of 18 countries pricing a separate list of products to link household consumption PPPs across regions, all countries priced a set of global core prices that were also imbedded in the regional lists. The result was that basic heading linking factors were based on all countries in 2011 instead of 18 countries subjectively selected in 2005. The concept of “importance” was introduced to ensure products making up a larger share of expenditures received more weight than products that may be available but not frequently purchased. A new method of aggregation coined the CAR method (Country Aggregation with Re-distribution) was used to aggregate PPPs to the GDP. The impact of the CAR method can be simulated; the impact was to raise the real expenditures in Asia 9 percent for 2005 and 6 percent for 2011. It is not possible to assess the use of the core list by all countries because the data required are not available for 2005.

---

<sup>1</sup> The six regions were Africa, Asia/Pacific, Commonwealth of Independent States, Eurostat/OECD, Latin America, and Western Asia.

<sup>2</sup> World Bank (2013), “Measuring the Real Size of the World Economy”

Housing rent PPPs in both benchmarks were to be either based on actual rentals or, if there was no rental market, then quantities of housing as measured by square meters by type of dwelling along with measures of quality. As a last resort, numbers of dwellings by type of dwelling were also used in some cases.

Housing rent PPPs in Africa and Asia were imputed in 2005 and again in Asia in 2011 because their Regional Coordinators considered that the data collected were not sufficiently consistent to produce reliable estimates of rent PPPs. The other regions used either rental prices or quantities adjusted for quality.

While essentially the same set of salaries data by occupation were obtained for the 2005 and 2011 rounds for government compensation PPPs, there were differences in which regions adjusted their data for productivity. In 2005, Africa, Asia, and Western Asia adjusted the compensation rates for productivity, but inter-regional linking factors were not adjusted. In 2011, the linking factors applied to all regions were adjusted for productivity, thus improving the comparability of government compensation.

Construction is one of the so-called “comparison-resistant” aggregates in the national accounts. A new method (based on pricing a combination of basic inputs, some composite inputs, wages and equipment hire) was developed in the 2005 ICP but it proved to have several shortcomings. As a result, the method was revised in 2011, with prices being collected for 38 material inputs, 7 categories of labor and the hire rates for 5 types of equipment. The TAG also recommended that data should be collected on mark-ups (profits, , project overheads, site costs etc.) and professional fees. However, many countries did not supply any mark-ups data and the data provided by the rest of the countries was in many cases of questionable quality, which meant that an unadjusted input approach had to be adopted in the 4 regions coordinated by the Global Office. For linking purposes, ten Eurostat-OECD countries provided data for the ICP method (input prices).

It is difficult to directly quantify the impact of the different methodologies. However, three issues probably contributed the most to differences between 2005 and 2011 outcomes. The use of the core vs. the ring greatly improved the quality of the linking process. Analysis suggests that adopting the global core list approach and the use of the importance classification would raise real expenditures in 2011 in Asia and Africa against the other regions, compared with the outcomes if the ring approach continued to be used. Analysis also points out that an adjustment for urban/rural prices in China would have raised its real expenditures in 2005 by as much as 10 percent. The CAR method in 2005 would have raised Asia’s real expenditures by about 9 percent compared with regions other than Africa.

The final results for 2011 will be of significantly better quality than the 2005 results because of the use of improved methodology. Not to be overlooked is the fact that the global office, regional coordinators, and country representatives had the benefit of experience gained from the previous round which was reflected in more thorough data validation and implementation of the more complicated estimation methods.

## 1. Introduction

The “2005 ICP Book” – “Measuring the real size of the world economy” by the World Bank (2013) – provides a comprehensive review of the statistical and economic theory underlying the estimation of Purchasing Power Parities (PPPs). Even though the PPPs provided by the ICP rest on a large body of statistical and economic theory, there are many decisions based on expert judgment that have to be made.

The outcome of each round of the ICP is a function of choices made, starting by determining what products to price and how to price them, choosing the index number formula to turn prices into basic heading PPPs, and then determining the multilateral formula to aggregate the PPPs to the GDP. Decisions about these different methods are made first at the regional level, then again for the process to link every country in every region to a numeraire currency.

The choices do not make much difference where the countries being compared have quite similar expenditure patterns and relative prices. However, when computing PPPs across countries such as Tajikistan vs. the US or Chad vs. the US the choices make a bigger difference in the results. And finally, there are aggregates of the GDP that are difficult to compare such as dwelling rentals, government expenditures, and construction that bring in another dimension of decisions to make.

Lessons learned from previous ICP rounds led to the development of several significantly new and improved methods for the 2005 round. Subsequent analysis of the 2005 data set the stage for additional improvements to the 2011 ICP.

The dilemma facing the ICP is that the continual improvement of methods is limiting the comparison of PPPs over time. While each benchmark may be based on the best methods available at the time, their comparability may be limited. Therefore, the purpose of this note is to describe the new methods implemented in the 2011 ICP, explain why they were chosen, and provide a subjective assessment of the potential impact of the change.

The complexity of the ICP and the choices to be made require expertise ranging from survey design to price and index number theory, the system of national accounts, and methods of aggregating PPPs to the GDP. For these reasons, a Technical Advisory Group was formed which comprises internationally known experts in each of these areas plus those who use the results for research, especially on poverty. The computational steps are intensive; therefore, a task force was formed to do the actual computations.

The following sections go into more detail about the choices made for ICP 2011 and how they affect comparability with ICP 2005.

## 2. Household final consumption expenditures: Selecting Products to be priced and classifying them as “important” or “less important”.

Statistical theory suggests that there should be a master frame listing every possible product purchased by consumers and the annual expenditures associated with each product for every country. A random sample of products would be selected for which national annual average prices would be determined. The expenditure on each product would be used to weight product PPPs to basic heading PPPs. The reality, however, is that there is no such list. While statistical theory can be used to determine the number of products to be priced, it is left to the regional and

national coordinators using their expert judgment to select the actual products out of the thousands of possibilities. The World Bank (2013) provides guidelines on the number of products to be priced; for example, it recommends that six products be priced for the rice basic heading compared to about 70 for the garment basic heading. Rice is a relatively homogeneous product, although it is necessary to specify different varieties to be priced, but garments are much more heterogeneous.

The comparability of the products being priced is an essential principle underlying the estimation of PPPs. A dilemma facing the ICP is that, while a product may be available in several countries; it may be a significant part of consumption in only a few countries. Since there are no data on expenditures for individual products, the relative prices or product PPPs are averaged with equal weights to obtain the basic heading PPP. To overcome this problem the Eurostat-OECD and CIS regions have adopted the concept of representativity to induce a form of weighting. A “representative” product is one that is purchased frequently by households and has a price level consistent with all products in the basic heading. Since “representative” products are those most frequently purchased, it is likely that they have lower price levels in countries where they are representative compared to countries where the product is available but not representative. Countries in the ICP regions attempted to use the “representative” classification in 2005 but were unable to consistently apply the notion of a “representative price level”. As a result, the concept was not used in 2005 in the ICP regions nor for estimating inter-regional linking factors.

The TAG proposed a simpler method for the 2011 ICP. Countries other than those in the Eurostat-OECD and CIS regions were asked to classify all goods and services for household consumption as either important or less important. Importance is defined by reference to the notional expenditure share of the item within a basic heading. The “importance” classification is a subjective process as was the assignment of “representative”, but simpler. If it is thought that the expenditure share would probably be large, the item is classified as important; if small, it is classified as less important.

The procedure to determine the products to be priced for household consumption in 2005 differed from what was done in 2011.

- In 2005, each regional coordinator in collaboration with national coordinators used Structured Product Descriptions (SPDs) furnished by the World Bank to create product specifications. Each region did so independently of other regions.
- After data collection and several iterations of data validation were completed, all regions submitted their final set of products that had been priced.
- The Global Office harmonized definitions and collapsed the combined lists from the regions into a list of about 1,000 products, coined the “ring list”. The ring list was the basis for a separate data collection by a subset of countries in each region which was then used to link PPPs across regions.
- The resulting price levels from the ring data collection were not consistent with those for the corresponding countries within each of the regions. For example, some ring countries priced products that were not representative of their consumption patterns, but received equal weight in the computation of linking factors. To the degree this took place, it points to an overestimate of price levels and an underestimate of real expenditures in Africa and Asia compared with the other regions.

Subsequent analysis of the ring list, prices, PPPs and linking factors provided several lessons learned. First the selected Ring countries did not always turn out to be representative of the other

countries in the region. Analysis showed that between-country variability was greater than the variability in relative prices within basic headings. Based on this analysis, the TAG recommended the following:

- Develop a set of core products that would be priced by all countries for linking purposes. The final 2005 Ring list became the starting point to determine the set of global core products for 2011.
- Include the core products in the regional lists as well. The starting point in each region to develop the list for 2011 was the 2005 regional product list and the set of global core products.
- Classify products as “important” or “less important”. Although countries were expected to be able to price a large number of core items, not all would have the same price levels or relative expenditures. Products common in some countries may be more difficult to find in other countries with the likelihood of higher prices. Therefore, the importance classification is needed to prevent an upward bias in the price levels used to estimate the between-region PPPs.
- Aggregate product PPPs to the basic heading using the Country Product Dummy-Weighted (CPDW) method with weights of 3:1 for important vs. less important, respectively.

Since the representative classification was only used in the Eurostat-OECD and CIS regions in ICP 2005, the likely result is an upward bias in PPPs (i.e. smaller real expenditures) in the remaining ICP regions. If the use of the importance classification were successful in 2011, the result would be lower relative prices and larger real expenditures in the remaining regions. These level differences are difficult to quantify because one would assume the regional and national coordinators, because of their previous experience, were able to better validate prices for the 2011 comparison.

### **3. Housing Rents**

Housing rent has proven to be one of the most problematical components in each ICP round, partly because the values are estimated so poorly in many countries’ national accounts and partly because the prices provided for the ICP in many cases are not consistent with the national accounts values. It is difficult to estimate PPPs for housing rents because of the varying mix of rental versus owner-occupied housing. PPPs for dwellings were computed three different ways in 2005. Rental rates were used where there was a large rental market. Where there was not a sufficient rental market, PPPs were computed indirectly using a quantity approach. The rental market in Africa and Asia was not large enough to use market rents to estimate PPPs, therefore the regions attempted to use the quantity method. However, this produced implausible results so PPPs were imputed for the African and Asia regions using the reference volume method. Even though there was not sufficient data to use the quantity method within those regions, there was enough to compute between-region linking factors.

- Because of the importance of housing PPPs, it was agreed by the TAG and the regions to place a greater emphasis on obtaining rental data. The TAG recommended that all countries provide two sets of data. First, all countries were to redouble efforts to provide rental prices. In addition, all countries were to provide data on quantities and quality even where there was a large rental market.

- Global specifications were prepared to collect data on rents and quantities which meant that the within-region PPPs were based on the same data to compute the between-region PPPs or linking factors

The dilemma is that countries without rental markets also had difficulty providing consistent and comparable quantity and quality data. For that reason, the Asia region is imputing PPPs for dwellings in the same way it did in 2005 (i.e. a reference volume method). Rental data were used to estimate dwelling PPPs in Africa, Western Asia, and Latin America, quantity data in the CIS region, and a combination of both in Eurostat-OECD. Regions will be linked using a combination of rental and quantity data for the subset of countries able to provide them.

While PPPs for dwellings are not optimal, the results between 2005 and 2011 are mostly comparable.

#### **4. Government Compensation**

For the 2005 ICP, the global office prepared a global list of 40+ government occupations for which countries provided annual salaries. Government salaries were adjusted for productivity in Asia, Africa, and Western Asia because of the huge differences between the wages paid in the countries in those regions. (Not adjusting for productivity differences would have resulted in some implausibly large estimates of government consumption expenditure in lower-income countries.) The annual salaries from the same list of occupations were used to compute between-region linking factors; however, they were not adjusted for productivity.

For ICP 2011, the list of global occupations remained about the same as those used in 2005. A major change is that the Eurostat-OECD region now uses output indicators to estimate real expenditures on education, while the other regions continue to use input indicators (wages). Therefore, it was necessary to develop some special procedures to link the Eurostat-OECD region to the other regions for the education PPPs. The occupation data for all regions were adjusted for productivity for the final linking thus improving the quality of the resulting PPPs and real expenditures.

#### **5. Construction**

Prior to the 2005 ICP round, PPPs for construction were based on an output (model-based) approach. The pricing methods for construction were changed in the 2005 ICP, partly for cost reasons (pricing models required specialist staff, such as quantity surveyors) and partly for methodological reasons (it had proven to be virtually impossible to specify a small number of models that were relevant to all countries in a region). The method adopted in the 2005 ICP was known as the “Basket of Construction Components” (BOCC) approach.

The BOCC approach involved collecting prices for a range of major construction components and basic inputs that were common across countries. The term “construction components” was used to describe specific physical outputs that are produced as intermediate steps in construction projects. A key element in this process was that the overall price estimated for each composite component related to an installed component including the costs of materials, labor and equipment meaning the price was more related to an output price rather than to an input price.

The objective of the BOCC approach was to provide simple and affordable price comparisons for construction. An important goal was to develop a technique that would enable construction to be priced in major locations within each country, which would result in comparable prices for

similar components across countries that had different labor/equipment mixes due to their different levels of economic development.

In practice, the BOCC method did not prove to be satisfactory. The main problems were the difficulty of pricing the composite components (they required construction specialists) and overlaps between the composite and the basic materials that also had to be priced to ensure adequate coverage of products.

The approach initially proposed for the 2011 ICP was based on pricing inputs (basic materials and equipment hire) and using them to approximate an output price on the basis of the relationships between the outputs and inputs for each country, as estimated in each country's input-output tables. Investigations showed that up-to-date input-output tables were not available in a sufficient number of countries for the approach to be viable. At its fifth meeting, in April 2011, the TAG endorsed the proposed (input) method and recommended the following:

- Basic heading PPPs for construction should be based on a simple combination of three groups of inputs (i.e. materials, labor and equipment) rather than allocating each input to model projects, or weighting each input in any other way.
- An unweighted CPD should be used to estimate PPPs within each of the three product groups for each basic heading. Each basic heading will have three PPPs—one each for materials, labor, and equipment. These will be combined using the respective weights.
- Basic Headings PPPs will be computed as weighted averages of the PPPs for each of materials, labor and equipment, with the weights being centrally determined for five clusters of countries in each region (although countries can provide their own specific weights, if available, rather than having a cluster-based weight applied). Note: This was later changed to three clusters (high, middle and low income).
- It may be necessary to adjust for different levels of labor productivity in the different groups of countries and some testing will be carried out by the Global Office.
- Prices are to be collected for 38 material inputs, 7 categories of labor and the hire rates for 5 types of equipment.
- Countries will be requested to confirm which resources are relevant for each basic heading.
- Information on mark-ups (profits, VAT, project overheads etc.), which will enable the PPPs for each basic heading to be adjusted to account for mark-ups, will be collected from construction experts. Ideally, mark-ups would be specific to each country but it may be necessary in some cases to estimate mark-ups for a group of similar countries.

At its seventh meeting in October 2011, the TAG discussed whether or not labor productivity adjustments should be applied. It concluded that there is no need to adjust for labor productivity differences between countries because each country had to provide weights for labor, materials and equipment hire for each basic heading. However, an assumption underlying not making this adjustment is that total factor productivity (TFP) is equal across countries.

At its eighth meeting, in May 2013, the TAG examined the construction data collected. The data quality was poor in a number of areas, particularly relating to “relevance indicators” (the types of materials being used in different types of construction) and the overheads for mark-ups and professional fees. In many cases, countries did not provide any estimates of these mark-ups. As a result, the TAG recommended that:

- A single set of relevance indicators is to be used within each region rather than those provided by individual countries. Each region will use a construction expert to provide advice on the relevance of the various components to the different types of construction activity.
- Resource weights provided by the Global Office will be used to average PPPs for materials, equipment, and labor for countries not able to provide the data.
- The construction prices should not be adjusted for mark-ups and professional fees, given the poor quality of the data collected on these aspects.
- The prices for equipment hire should be split into those including an operator and those excluding an operator and treated as separate product specifications.

In September 2013, at its ninth meeting, the TAG reconsidered its earlier recommendation that no productivity adjustment be applied to construction labor and the associated implication that TFP in construction is identical across countries. The TAG considered the possibilities for specific adjustments for labor productivity or for TFP. The consensus of the TAG was that no adjustment should be applied for labor productivity but that an adjustment for TFP should be considered. A set of TFP adjustment factors was produced but it was clear that they added noise to the construction price estimates rather than improving them. As a result, the TAG reaffirmed its earlier decision that no adjustments should be applied to construction prices for either labor productivity or TFP differences across countries.

Fully implementing the TAG's recommendations regarding adjustments to prices for construction markups proved problematical because of the poor quality (or non-provision) of data by participating countries. Data are available for some countries but they have not been properly validated because of the Regional Coordinators' views that sufficient countries had not provided data on markups and there was a large degree of variability in the data that were available. At this late stage, it is not possible to ask countries to collect new markup data for 2011.

Because of problems in 2005 pricing the composite components, the construction PPPs were essentially based on the basic components which meant the resulting PPPs mostly reflected an input approach. The basic components priced in 2011 were about the same as those priced in 2005. The net result is that neither 2005 nor 2011 PPPs were based on output prices. While this is not the desired method, it can be said that the construction results for 2005 and 2011 are broadly comparable.

## **6. Estimating within-region PPPs.**

Chapters 4<sup>3</sup> and 5<sup>4</sup> in "Measuring the Real Size of the World Economy" describe the different properties of the various indexes that can be used to compute basic heading PPPs and aggregate them to the GDP. These chapters are also the basis for the choices made for the 2011 ICP. The basic methodology used in 2005 follows.

- The Country Product Dummy method (CPD) was used in the ICP regions and the Jevons-GEKS\* used in the Eurostat-OECD and CIS regions to compute basic heading PPPs. The Jevons-GEKS\* method made use of the representative classification, the CPD did not.

---

<sup>3</sup> Chapter 4: Computation of Basic Heading PPPs for Comparisons within and between countries

<sup>4</sup> Chapter 5: Methods of Aggregation above the Basic Heading Level within-regions

- The GEKS method was used to aggregate basic heading PPPs to the GDP in all regions except Africa where the Iklé method was used. Both provide results that are transitive and base country invariant. The base country invariant property ensures that the PPPs between any two countries are the same no matter which country is the base. The transitive property simply means that the price level, for example, of the UK relative to the US is the same whether it is calculated directly or through any possible chain of countries—US to Nepal to Nigeria, etc. By imposing the transitive property, the PPPs between any two countries can change if the mix of the remaining countries changes. When a region is homogeneous, the direct and indirect PPPs remain similar. However, the process induces more variability when indirect PPPs enter from countries with widely different price and economic structures. This has implications for the linking methods discussed in the next section. The Iklé method provides additive results but can also overestimate price levels in poorer countries.
- Dwelling rent PPPs in Africa and Asia were imputed, with the other regions using either rental prices or quantities adjusted for quality. Government salaries were adjusted for productivity in Asia, Africa, and Western Asia. While these steps improved the results within the Asia and Africa regions, they affected their comparability with countries in other regions. This especially applied to government compensation because the between-region PPPs were not adjusted for productivity.
- Prices for each product were to represent national annual average prices. Where this was not possible, countries were to use CPI and other information to calibrate the prices to national annual averages. There was much debate about China which submitted prices that mostly represented urban areas. Experts such as Deaton (2008)<sup>5</sup> estimated that adjustments based on the distribution of consumption between urban and rural areas would raise the estimates of real Chinese GDP by about 10 percent in 2005.

The methodology for 2011 per TAG recommendations follows:

- The CPD-W for household consumption was used in ICP regions with the weights of 3:1 for important vs. less important products. The Jevons-GEKS\* as used in Eurostat-OECD and the CIS in 2005 remains the same.
- The GEKS was used to aggregate basic heading PPPs to the GDP – the same as used in 2005.

The issue is whether the relative rankings within each region are comparable between 2011 and 2005. The within-region 2011 results will not be exactly comparable with those for 2005 if the methods used to estimate dwelling rents and the application of productivity adjustments differ between the two periods. The set of countries changed between some regions with Chile for example becoming part of the OECD, major countries such as Argentina not taking part, and others such as Iran treated separately. In other words, changes in the relative rankings of countries within-regions can be the result of changes in methodology.

## **7. Linking the regions**

The methodology to link the regions remained about the same between 2005 and 2011 at the basic heading level except that all countries provided prices for a set of global core products

---

<sup>5</sup> Deaton, A., and A.H. Heston. (2008) “Understanding PPPs and PPP-based National Accounts” NBER Working Paper No. 14499, National Bureau of Economic Research, Cambridge, MA

instead of 18 countries providing prices for the separate ring list. However, a significant change was made in how the linking takes place for higher aggregates up to the GDP.

A review of the linking method at the basic heading level and resulting issues follows:

- Core product prices provided by all countries (18 ring countries in 2005) are deflated to a regional currency using within-region basic heading PPPs
- The result is five sets of regional prices treated as “super countries”. The CPD-W (CPD in 2005) over these five sets of regional prices provides between-region basic heading PPPs. These between-region PPPs when multiplied by within-region basic heading PPPs convert to a global currency. The same regional scalar, say for rice, times each country’s within-region PPP converts it to a global PPP.
- This method preserves within-region fixity which means the relative rankings between countries in the same region remain the same after linking.
- In 2005, this computation step only included the 18 ring countries and the ring prices whereas in 2011 the between-region PPPs were based on core prices provided by every country. Since these core prices were also included in the estimation of within-region PPPs, the between-region results were more consistent with the within-region results.
- The CIS region in both 2005 and 2011 was linked to the Eurostat-OECD region using Russia as a bridge country.

Between-region linking factors for 2011 are based on all countries instead of a subset subjectively selected; thus are more statistically robust. The final computational step was to link regions at higher level aggregates and the GDP. In 2005, the between-region PPPs (linking factors) were aggregated to the GDP using the GEKS. In a separate computation, within-region PPPs were aggregated to the GDP as described in the previous section. Again, the aggregated between-region PPPs times the aggregated within-region PPPs calibrated the results to the global currency. Chapter 6 in “Measuring the Real Size of the World Economy”<sup>6</sup> reviews the properties of this method which showed that the computations are dependent on the choice of base country. As a result, the TAG, after considering several alternatives, proposed that a method called the CAR (Country Aggregation with Re-distribution) be used.

A review of this method and resulting issues follows:

- A global aggregation that includes all 150+ countries and 155 Basic Headings in a GEKS aggregation provides PPPs calibrated to a global currency. In order to preserve within-region fixity, real expenditures are summed to regional totals which are then distributed within each region according to the distribution from the within-region computations. These results are base country invariant, transitive, and preserve fixity.
- The global PPP between any two countries is the geo mean of the direct comparison and the n-2 indirect comparisons with every other country. The range of the direct and indirect comparisons is small for countries with similar economic and price structures, but can become large where countries differ significantly.
- Simulations<sup>7</sup> show that real expenditures are increased in Asia by 9 and 6 percent respectively for 2005 and 2011. This simulation does not include the impact of using global core prices from all countries vs. ring prices from 18 countries.

---

<sup>6</sup> Chapter 6: “Methods of Aggregation above the Basic Heading Level: Linking the Regions”

<sup>7</sup> ICP 2011 Special Task Force Report to the Executive Board. December 26, 2013

## 8. Summary

The ICP includes economies ranging from city-states and small islands to large and diverse countries such as Brazil, Russia, India, and China. Like all statistical endeavors, PPPs are statistical estimates that fall within some margin of error of the unknown, true values. The ICP Global report<sup>8</sup> suggests using caution when comparing economies by the size of their GDP or in per capita expenditures because there may be errors in the calculation of GDP and population sizes in addition to the statistical variability inherent in the PPPs. The 2005 global report indicated that differences in GDP less than 5 percent lie within the margin of error of the PPP estimation. Deaton (2013),<sup>9</sup> suggested a method to measure statistical variability in the estimation of PPPs which comes from the choice of products, the range of PPP product prices, and differences in basic heading expenditure weights. He showed that standard errors of PPPs become larger for countries with different price and economic structures with approximations showing that the standard errors of Indian or Chinese prices to American 2005 prices are between 10-15 percent.

The range suggested by the standard errors reflects the variability resulting in choice of methods. The analysis above points out that a possible adjustment for urban/rural prices in China would have raised its real expenditures by 10 percent in 2005. The CAR method would have raised real expenditures for all Asian countries in 2005 by about 9 percent compared with the other regions (excluding Africa, whose real expenditures would have also been increased by using the CAR method). Therefore, adopting the CAR approach for 2011 rather than the ring approach would raise all Asian countries' real expenditures in 2011 compared with Eurostat-OECD countries.

---

<sup>8</sup> Global Purchasing Power Parities and Real Expenditures - 2005 International Comparison Program

<sup>9</sup> Deaton (2012). "Calibrating measurement uncertainty in purchasing power exchange rates" Paper presented at the 7<sup>th</sup> meeting of the ICP Technical Advisory Group, September 17-18, 2013, Washington, DC