Quality Assurance Procedures for Estimating Purchasing Power Parities

Marko Olavi Rissanen[1], Nada Hamadeh[2]

[1] *Statistician; The World Bank Group, Washington, DC, USA; mrissanen@worldbank.org*
[2] *Senior Statistician; The World Bank Group, Washington, DC, USA; nhamadeh@worldbank.org*

Abstract
The International Comparison Program (ICP) is a worldwide statistical initiative designed to estimate purchasing power parities (PPPs). PPPs are conversion rates that convert to a common currency while equalizing the purchasing power of different currencies, hence allowing for accurate volume measures of economic activity as well as comparisons of standards of living across countries. Among many other uses, the Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs) rely on PPP-based indicators for their key goals on economic growth, poverty, energy and environment. To estimate PPPs for the world's economies, the ICP conducts surveys to collect price and expenditure data for all goods and services that make up the gross domestic product (GPD). Significant resources are allocated to data collection and validation in order to ensure PPP estimates of the highest possible quality. Validation procedures are carried out first at the national level to ensure intra-country data quality, subsequently at the regional level to ensure inter-country data quality, and finally at the global level to ensure inter-country data quality across the regions. Moreover, efforts are devoted to validating the resulting regional and global PPPs from an economic point of view to ensure their soundness. The quality procedures require close cooperation and collaboration between participating countries, regional implementing agencies and the global implementing agency. Frequent meetings, training sessions and validation workshops are organized to ensure consistency of methodology, comparability of prices and soundness of resulting PPPs. The ICP Book (World Bank, 2013) and the ICP Operational Guidelines (World Bank, 2015) were also developed as reference materials to ensure a common understanding of PPPs and their underlying methodology, quality assurance and data validation procedures.

Keywords: International Comparison Program (ICP), Purchasing Power Parities (PPPs), Sustainable Development goals (SDGs).

1. Introduction

1.1. The International Comparison Program

The International Comparison Program (ICP) is a worldwide statistical initiative designed to estimate purchasing power parities (PPPs). PPPs are conversion rates that convert to a common currency while equalizing the purchasing power of different currencies, hence allowing for accurate volume measures of economic activity, as well as comparisons of standards of living, across countries. Among many other uses, the Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs) rely on PPP-based indicators for their key goals on economic growth, poverty, energy and environment.

In order to estimate PPPs for the world's economies, the ICP conducts price surveys and collects national accounts expenditure data for all goods and services that make up the gross domestic product (GPD). To date, the global ICP rounds have been conducted every six years, 2011 being the most recent round with nearly 200 participating countries (World Bank, 2014). In March 2016, the United Nations Statistical Commission (UNSC) recommended in its 47th Session to conduct the next ICP round for the reference year 2017, and to carry out subsequent rounds at more frequent intervals (United Nations, 2016).

A key difference between production of typical national statistics, such as the consumer price indices, and PPPs is that the latter is a multilateral index, containing data from all participating countries. As a result, the production of PPPs and related quality assurance procedures, entail close cooperation and collaboration between the *national implementing agencies*—typically national statistical offices of the participating countries—, *regional implementing agencies*—multilateral development banks and regional agencies—, and the *global implementing agency,* whose role has been assumed by the World Bank. The 2011 round of ICP covered eight regions, seven of them geographical: Africa, Asia, Commonwealth of Independent States (CIS), Latin America, the Caribbean, Western Asia, and the Pacific Islands. The eighth region comprised the economies participating in the regular PPP program coordinated by the Statistical Office of the European Communities (Eurostat) and the Organisation for Economic Co-operation and Development (OECD).

1.2. Overview of PPP estimation methods and quality assurances processes

The PPPs estimation process starts with the national implementing agencies of the countries participating in the regional comparisons providing the regional implementing agencies with a set of monthly, quarterly or annual national average purchasers' prices for items chosen from a common list of precisely defined items. These common lists include both regional items, priced in the given region, as well as global items, priced in all ICP regions. These sets of prices cover the whole range of final goods and services included in the GDP: household consumption expenditures; government expenditures; and gross fixed capital formation expenditures.

At regional level, these prices are used to calculate regional PPPs for basic headings. Basic headings are the lowest aggregation level in the ICP national accounts expenditures classification for which explicit expenditure weights can be estimated. A common method to calculate basic heading PPPs is the country product dummy (CPD) method, as well as its variants, such as the weighted-CPD method (World Bank, 2013, 2015). The CPD method was first introduced by Summers (1973), and later revisited, among others, by Diewert (2005) and Rao (2009). The regression equation for the CPD can be written as:

$$\ln p_{cp} = y_{cp} = x_{cp}\beta + \varepsilon_{cp} \qquad (1)$$

$$x_{cp} = \left[Dc_2 \ldots Dc_{Nc} Dp_1 Dp_2 \ldots Dp_{Np} \right]$$
$$\beta = \left[\alpha_2 \ldots \alpha_{Nc} \gamma_1 \gamma_2 \ldots \gamma_{Np} \right]^T \qquad (2)$$

where p_{cp} is the price of product p in country c; Dc_j and Dp_i are country and product dummies, respectively; and Np and Nc are number of products and countries, respectively.

The regional basic heading PPPs are further combined with national account expenditures and aggregated using aggregation methods, the most common being the Gini-Éltetö-Köves-Szulc (GEKS) method, in order to derive intraregional measures of price and volume relatives for the countries participating in the regional comparison (World Bank, 2013, 2015). The GEKS method was first introduced by Gini (1924, 1931), and later independently rediscovered by Éltetö and Köves (1964) and Szulc (1964). This non-additive method is based on Fisher's

bilateral indexes (Fisher, 1922) and it consists of two calculation stages. First, the basic heading PPPs are aggregated using the national accounts expenditure weights to obtain the bilateral Fisher PPPs (F-PPPs) for all pairs of countries. For any two economies j and k, the binary F-PPP are computed as:

$$P_{F\ j,k} = \left(\left(\frac{\sum_i p^i_j q^i_k}{\sum_i p^i_k q^i_k} \right) \left(\frac{\sum_i p^i_j q^i_j}{\sum_i p^i_k q^i_j} \right) \right)^{1/2} \quad (3)$$

where p^i_j is the basic heading PPP for basic heading i in country j, and q^i_j the respective expenditure weight for basic heading i in country j.

Second, all direct and indirect F-PPPs are averaged geometrically to obtain the transitive GEKS-PPPs. The GEKS-PPPs are computed as:

$$P_{GEKS\ j,k} = \left(\prod_l P_{F\ j,l} \Big/ P_{F\ k,l} \right)^{1/K} \quad (4)$$

where K is the total number of economies.

At global level, the regional PPPs are linked to form a global set of PPPs and related measures of price and volume relatives. In order to link the various regional basic heading PPPs, the so-called inter-regional linking factors are calculated based on prices of global items collected in all regions using the CPD method. The resulting linked basic heading PPPs are subsequently weighted and aggregated using the GEKS method, and finally adjusted to maintain the fixity of the regional results, based on a country aggregation with volume redistribution, or CAR-volume, procedure (World Bank, 2015). The CAR-volume procedure redistributes the total regional expenditure volumes from the global aggregation according to the countries' expenditure volumes shares in the regional comparisons. This allows subsequently calculating the global PPPs indirectly, that is, by dividing countries nominal expenditure volumes by the real expenditure volumes. As noted above, the CAR procedure enables maintaining fixity between the global and regional PPPs.

Significant resources are allocated to data collection and validation in order to ensure PPP estimates of the highest possible quality. Validation procedures, which have been developed

throughout the various ICP rounds, are carried out first at the national level to ensure intra-country data quality, subsequently at the regional level to ensure inter-country data quality, and finally at the global level to ensure inter-country data quality across regions. Moreover, efforts are devoted to validating the resulting regional and global PPPs from an economic point of view to ensure their soundness. Frequent meetings, training sessions and validation workshops are organized to ensure consistency of methodology, comparability of prices and soundness of resulting PPPs. The ICP Book (World Bank, 2013) and the ICP Operational Guidelines (World Bank, 2015) have also been developed as reference materials to ensure a common understanding of PPPs and their underlying methodology, quality assurance and data validation procedures. Overall, the ICP quality assurance methods and processes stem from those developed within the Eurostat-OECD PPP program (Eurostat and OECD, 2012).

1.3. Quality assurance objectives

As explained above, basic heading PPPs, or in other words elementary PPPs, are calculated based on prices collected by national implementing agencies for a common list of precisely defined items. These elementary PPPs are subsequently weighted with national accounts expenditures and aggregated to calculate PPPs for the GPD and its components, such as household consumption and gross fixed capital formation. The following sections of this paper cover the quality assurance phases during the intra-country, inter-country and global validation stages separately; however, in general, price validation processes have two main objectives.

The first objective is to clear data from pricing errors. Two types of non-sampling pricing errors can be identified, namely, *price error* and *product error*. A price error occurs when price collectors price products that match the item specification, but they fail to record the price, or related quantity or unit of measure, correctly. A price error can also occur during the process of reporting and transmitting a price that was initially recorded correctly. A product error occurs when price collectors price products that do not match the item specification and neglect to report having done so.

The second objective is to ensure comparability of prices. The purpose of the price comparison is to compare comparable products, or in other words "likes with likes."

Incomplete, unclear, or too loose item specifications may introduce product errors because the products whose prices are being compared may not be truly comparable. For example, if a generic item specification—that is, a specification that does not define the brand and model to be priced—is used, it may embrace a wider range of products than was intended when the specification was created. In this case, errors may be introduced into the estimated PPPs because products priced between countries differ significantly in their price determining characteristics, such as brand and model. The role of the regional and global implementing agencies is to ensure that collected prices are comparable within the region and across the regions, by analyzing both the price data and resulting PPPs. Because product errors at the regional or global levels may not be errors *per se* at the country level, close collaboration between the implementing agencies is a necessity.

Overall, it should be stressed that prevention is preferable to correction; the incidence of non-sampling errors can be significantly reduced through good survey design and management. Price collections must be carefully planned, efficiently carried out, and properly supervised. Item specifications must also be sufficiently detailed to enable price collectors to identify products unambiguously in the outlets they visit. In other words, validation procedures complement good survey practices.

2. Intra-country quality assurance measures

Intra-country validation is the first stage of the ICP quality assurance processes. Within this stage, the objective is to verify that price collectors within the same country have priced products that match the item specifications and that the prices they have reported are correct. These checks are carried out without reference to the price data of other countries. The intra-country validation stage consists of three validation phases, namely, *initial data validation*, *statistical tests*, and *finalization of price data*. During the first phase, the price data generated by price collectors are merged and the first checks are made. These checks include ensuring the validity of the item codes, observed quantities and units of measurement, sub-national and temporal price variations, and several other details.

For the first phase, the national implementing agencies use their local knowledge to determine the validity of the sub-national price data, that is, price data collection from different regions within a country. The relationship of the various areas to the country as a whole resembles the one between the countries and the ICP region; even large differences in prices among different parts of the country may be appropriate, however, representative national average prices have to be the end result. The temporal validation checks are possible in cases where prices for the same items are available either for different quarters within a reference year, or between the ICP rounds. The challenge with the latter however is the multi-year gap between the ICP rounds, and relating to this, possible item specification modifications or significant market changes.

During the second phase, data are analyzed more systematically using four statistical tests or measures: coefficient of variation (CV), maximum/minimum price ratio, standard deviation, and t-value. The detailed price information and the statistical tests are examined using two tables, namely, the *average price table*, which shows average prices as well as CVs, and maximum/minimum price ratios for each item, and the *price observation table*, which gives the standard deviations and t-values for each individual price observation. Editing for product errors and price errors involves identifying prices whose statistical test values are determined to be either too high or too low vis-à-vis the average according to given criteria. The price may score a value for a given test that exceeds a pre-determined limit, or its value may fall outside some pre-specified range of acceptable values. Both are typical ways of detecting errors in survey data, and both are employed by the ICP.

It is important to note that the prices with problematic values are not necessarily wrong, however, the fact that their values are considered to be outside a pre-specified range suggests that they could be wrong, and hence they need to be investigated. It is not an ICP practice to automatically reject prices with problematic values, but instead flag them for review. Prices with high values that are found to be wrong are errors and should be corrected or dropped, whereas prices with problematic values that are shown to be accurate observations are

"outliers" and should be retained. However, in practice, some extreme outliers may need to be discarded, in order to remove the "noise" they introduce into the data.

During the final phase, the validated average prices are prepared for submission to the regional implementing agencies, accompanied with quality indicators such as their CVs, maximum/minimum price ratios, and number of underlying price observations.

3. Inter-country quality assurance measures

Inter-country validation is a collective validation stage involving the regional and national implementing agencies, and it is designed to ensure that collected prices are comparable and that item specifications have been interpreted in the same way across countries. The inter-country validation is an iterative process requiring a number of rounds of editing and verification. Possible errors are uncovered by identifying prices that diverge significantly from other prices in the dataset. When errors are identified and corrected, the parameters of the price dataset will change, and, subsequently, divergence measures of each price remaining in the dataset will change. A second edit will find new possible errors that will need to be verified. Again, when actual errors are corrected, the parameters of the price dataset will change, which may lead to more possible errors being detected if a third edit is done. After each round, as incorrect or incomparable prices are removed or corrected, the PPPs will become more reliable.

The inter-country validation stage consists of several phases. The first phase, *initial data validation,* is a quick validation phase to reveal any rough or extreme problems with the data. *Validation at the basic heading level* is the main phase of the inter-country validation, and most of the edits are carried out within this phase. *Validation at aggregated levels* take average prices to a wider context, above the basic headings. *Temporal analysis* checks the consistency of PPPs with regard to a previous ICP round or previous quarters within an ICP round. *Finalization of data* concludes the inter-country validation phase.

The *initial data validation* phase is a necessity because, when the data received from the national implementing agencies are compared for the first time, a number of severe errors are usually evident, making any in-depth analysis meaningless. The most extreme errors are usually unit of measurement or quantity-related, or simple typing mistakes that may not be visible at the country level, but become clear when data for different countries are being compared.

As noted above, *validation carried out at the basic heading level* is the main phase of the inter-country validation. The key tools for this work are the Quaranta table, named after Vincenzo Quaranta, who first proposed it for use in the European PPP program in 1990, and the Dikhanov table, named after Yuri Dikhanov, who first proposed it for use during the 2005 ICP round (World Bank, 2015). The purpose of both validation tables is to screen the national average prices for possible errors by comparing the average prices of the same items across countries. Both tables provide similar measures of average price variation for basic headings, countries, and items. For the comparison of average prices, the tables provide three main measures: exchange rate-ratios, PPP-ratios (CPD residuals in the Dikhanov table), and price level indices (PLIs). For analysis of price variation, the tables provide four variation measures, ranging from variation of underlying price observations to variation of average prices of items within a basic heading.

For the comparison of average prices, the prices, reported in national currencies, need to be first converted to a common currency. However, prices cannot be compared across items directly, even when expressed in the same currency. Instead, the price ratios of an item can be compared with the equivalent price ratios for other items, once they have been "standardized." The standardized price ratios for an item are the ratios of the individual country average prices to the geometric mean of the average prices in all countries pricing the item, when the average prices are expressed in a common currency. Both exchange rates and PPPs are used to convert the average prices to a common currency, and both the exchange rate-converted average prices and the PPP-converted average prices are used to derive the standardized price ratios. The standardized price ratios based on exchange rate–converted prices are called *XR-ratios*, and

the standardized price ratios based on PPP-converted prices are called *PPP-ratios*. A high (low) XR-ratio or PPP-ratio means that the national average price for the item in question is high (low) compared with the average prices of the same item in other countries, and there comes a point at which the XR-ratio or PPP-ratio is too high (low) to raise the question of whether the average price may be erroneous.

An XR-ratio that lies outside predetermined limits may signal a questionable average price. However, the principal reason for calculating PPPs is that when the prices of a given item are converted into a common currency using exchange rates, they are not in fact equal in all countries. The general level of prices tends to be systematically higher or lower in some countries than in others. Thus a high (low) XR-ratio for an individual item in one country may be largely due to the fact that the general price level for that country is high or low when exchange rates are used, and hence it may not signal any abnormality in that particular average price. For this reason, XR-ratios are less useful than PPP-ratios for validation purposes.

On the other hand, PPPs are the rates of currency conversion that are designed to equalize price levels for the items covered. The PPPs for a basic heading are the rates of currency conversion that should enable a given amount of currency to purchase the same basket in all countries. Thus a high (low) PPP-ratio for an individual item in one country signals that the price is having a different relative position among the other items the country has priced for the basic heading. The price may thus be either questionable, requiring further review, or the item may be atypical for some reason among the others items within the basic heading for the country in question. In either case, high (low) PPP-ratios are flagged for further investigation and are subject to either verification or correction.

The third measure used to compare prices is the *price level index* (PLI). This index is defined as the ratio of the basic heading PPP to the exchange rate and is expressed as a percentage. The PLI for the base – country or region – is equal to 100. A PLI greater (less) than 100 indicates that when the national average prices are converted at exchange rates, the resulting prices within the basic heading tend to be higher (lower) on average than the prices in the base country or region.

Furthermore, as noted above, the validation tables provide four coefficients of variation. The first one is the *overall average coefficient of variation*, which measures dispersion among all the PPP-ratios for a basic heading. In doing so, it measures the homogeneity of the price structures of the countries covered by the basic heading, and the reliability of the PPPs calculated for the basic heading. Naturally, the higher the value of the coefficient, the less homogeneous the price structures are, resulting in less reliable PPPs.

The second measure is the *country coefficient of variation*, which measures dispersion among a country's PPP-ratios for a basic heading. In other words, it measures the variation in a country's PPP-ratios among the items in a basic heading, and the reliability of its PPP for the basic heading. The higher the coefficient's value, the less uniform the country's PPP-ratios, leading to less reliable PPPs.

The third measure is the *item coefficient of variation*, which measures dispersion among the PPP-ratios for an item. It is an indicator of comparability and accuracy that addresses the question of whether comparable products have been priced across countries. The higher the coefficient's value, the less uniform the item's PPP-ratios. Low uniformity raises questions about the comparability and accuracy of the item's pricing across countries.

The fourth measure is the *price observation coefficient of variation*, which measures variation in the price observations, on which the average price reported for an item by a country is based. It is taken straight from the average price table and is used to identify extreme values among average prices during the intra-country validation.

Besides serving as editing tools, the coefficients provide a means of monitoring progress during the validation stage and, at its conclusion, of assessing the effectiveness of the entire process of editing and verification in reducing the incidence of non-sampling errors among the price data. Coefficients should be significantly smaller at the end of validation than they were at the beginning.

After the validation at the basic heading level, the next phase is validation at the *aggregate level*. This phase aims at identifying inconsistencies that would not be potentially found when working at the basic heading level. These checks are made possible by running the validation tables at levels above the basic headings, that is, by combining items belonging to different basic headings under a single heading. The key objective of this phase is to identify possible errors that are not visible at the basic heading level. For example, if a country has reported prices systematically in a wrong unit of measurement for all items within a basic heading, this may be difficult to identify at basic heading level, however, the error becomes obvious when these prices are directly compared across items from other basic headings.

Temporal validation allows comparisons to be made using PLIs or PPPs over time. Comparisons of PLIs or PPPs over time is more robust than comparing the average prices for certain items directly, since item specifications may have changed between the two comparison points. Nevertheless, the basket of items as well as participating countries may have an impact on the results that is not related to the actual price level changes within individual economies. In any case, significant differences in countries PLI rankings are not expected between the comparison rounds.

Gradually, after a number of rounds of verification, a convergence will occur, and the return on further rounds of verification will be deemed marginal therefore not worth pursuing. Consequently, the inter-country validation stage will be concluded.

4. Global quality assurance measures

After the data are validated at the country and regional levels, the final validation stage is the global validation in which the regional and global implementing agencies verify the quality of regional PPPs, prices collected for global prices and subsequent inter-regional linking factors, and overall soundness of the global PPPs. As the regional validation stage, the global validation stage is conducted through an iterative process, consisting of several rounds of review and editing.

The overall objective of the first phase, verification of regional PPPs, is to ensure that the quality thresholds of the various regional comparisons are similar, and that there are no pending verification issues. This phase is conducted through re-generating and reviewing the regional validation tables, and checks covered in the previous section of this paper.

Subsequently, ensuring the comparability of prices collected for the global items across the regions, and the quality of the inter-regional linking factors, is the main phase of the global validation. This phase has three main steps. First, an assessment is needed of how well the global items reflect the regional results as a whole. This assessment can be carried out by analyzing the regional validation tables but, most importantly, by comparing regional PLIs and PPPs calculated based on a full set of priced items, solely based on global items, and solely based on regional items. The results of the three different scenarios are expected to have differences, but nevertheless they should yield broadly similar results.

Secondly, the price data for the global items are validated across regions by calculating PPPs and PLIs for all ICP countries. This step is challenging because of the large number of countries, making it difficult to apply the conventional validation tables. Hence, a top-down validation approach is applied, by analyzing the PLI and country CV matrixes, and moving to more detailed levels if problematic cases are encountered. The key means to further analyze the identified problematic cases are similar to the one described in the inter-country validation stage, that is, the review of XR- and PPP-ratios, and related coefficients of variation.

Thirdly, the inter-regional linking factors, calculated based on regional global price data, need to be analyzed and verified. Overall, the linking factors are expected to vary from basic heading to another, but they should be fairly stable among similar basic headings, such as those for food items, products items or services.

Moreover, the resulting final aggregated global PPPs need be further analyzed for possible discrepancies, derived either from the regional results, from the inter-regional linking factors, or from the global aggregation. For this, the so-called Paasche-Laspeyres spreads can be reviewed. As noted above, the GEKS-aggregated PPPs are based on the Fisher index, which is

the geometric mean of the Paasche and Laspeyres indexes. For each pair of countries (j and k) the Laspeyres index involves averaging basic heading PPPs by using country j's basic heading weights, while the Paasche index involves averaging basic heading PPPs using country k's basic heading weights. The degree of similarity of the Paasche and Laspeyres results for each bilateral PPP is measured by the relative difference between the two indexes. Hence, for a pair of economies with a large Paasche-Laspeyres spread, a review of the distributions of basic heading PPPs and expenditure shares is beneficial as one or both may include outliers deserving additional inspection.

Lastly, more general economic analysis can be applied for the review and validation of global PPPs, such as an assessment on how well the so-called "Penn-effect" (Samuelson, 1994), which is the observation that consumer price levels in richer countries are systematically higher than in poorer ones, and "Balassa–Samuelson effect" (Balassa, 1964), which is the assumption that productivity varies more by country in the traded goods' sectors than in other sectors, hold, especially vis-à-vis previous comparisons, or rather, extrapolations from previous comparisons. It should be, however, noted that all extrapolations for a long time span are known to suffer from various limitations (World Bank, 2014).

5. Conclusion

The quality assurance procedures and practices developed during the preceding ICP rounds have significantly improved the PPP estimation processes. The developed reference materials, including the ICP Book and Operation Guidelines, have enhanced the understanding of PPPs and their underlying methodology, quality assurance and data validation procedures, and in doing so increased the transparency of PPP estimation.

Going forward, it foreseen that the UNSC recommendation to conduct the ICP rounds more frequently will create avenues for further standardization and automation of the validation procedures and practices, especially through relying on web-based communication and data editing means.

References

Balassa, B. (1964), The Purchasing Power Parity Doctrine: A Reappraisal, Journal of Political Economy, 72 (6), pp. 584–596.

Diewert, W. E. (2005), Weighted Country Product Dummy Variable Regressions and Index Number Formulae, Review of Income and Wealth, 51 (4), pp. 561–570.

Éltetö, O. and Köves, P. (1964), On a Problem of Index Number Computation Relating to International Comparison, Statisztikai Szemle, 42, pp. 507–518.

Eurostat and OECD (2012), Eurostat-OECD Methodological Manual on Purchasing Power Parities, Eurostat Methodologies and Working Papers, Luxembourg.

Fisher, I. (1922), The Making of Index Numbers, Houghton Mifflin, Boston.

Gini, C. (1924), Quelques considérations au sujet de la construction des nombres indices des prix et des questions analogues, Metron, 4 (1), pp. 3–162.

Gini, C. (1931), On the Circular Test of Index Numbers, Metron, 9 (9), pp. 3–24.

Geary, R. G. (1958), A Note on Comparisons of Exchange Rates and Purchasing Power between Countries, Journal of the Royal Statistical Society Series A, 121 (1), pp. 97–99.

Rao, D. S. P. (2009), Generalised Eltetö-Köves-Szulc (EKS) and Country-Product-Dummy (CPD) Methods for International Comparisons, On Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications, ed. Rao, D. S. P., pp. 86–120, Edward Elgar, Cheltenham, U.K.

Samuelson, P. A. (1994), Facets of Balassa-Samuelson Thirty Years Later, Review of International Economics, 2 (3), pp. 201–26.

Summers, R. (1973), International Price Comparisons Using Incomplete Data, Review of Income and Wealth, 19 (1), pp. 1–6.

United Nationals Statistical Commission (2016), Report on the Forty-Seventh Session, United Nations, New York.

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

World Bank (2014), Purchasing Power Parities and the Real Size of World Economies: A Comprehensive Report of the 2011 International Comparison Program, World Bank, Washington, DC.

World Bank (2015), Operational Guidelines and Procedures for Measuring the Real Size of the World Economy: 2011 International Comparison Program, World Bank, Washington, DC.