Using Purchasing Power Parities to Compare Countries: Strengths and Shortcomings

Patrick Honohan
December 2020

Good international economic policy requires good comparative data on national economic performance. In May 2020, the World Bank published the latest comprehensive update of purchasing power parities (PPPs), as part of the International Comparison Program (ICP). Produced for 176 countries, these PPPs are used to convert data (especially GDP and other quantities based on national income accounts) from national currencies to a common basis in a way that is meaningful for economic analysis. The new PPP data are an invaluable tool for international economic comparisons, but they should be used with caution and an understanding of their limitations.

Using market exchange rates instead of PPPs can be highly misleading, especially in volatile economies, when market exchange rates fluctuate sharply and decouple from relative prices. PPPs are like synthetic exchange rates: If they prevailed in the market, prices would on average be the same across countries.

The collection of an increasingly extensive set of data on prices is a huge international statistical effort. Previous updates were accompanied by controversy. The estimated prosperity of countries like India and China was sharply revised, for example—downward in 2005 and then upward in 2011. The current update (which refers to 2017) has been less controversial. Although methodological debates continue, the current approach appears to be generally considered a reasonable compromise.

One of the signature uses of these numbers is to compare average living standards and the incidence of poverty based on an international standard, for which PPPs are indispensable. PPPs are not always appropriate, however. There has been a growing tendency to look at aggregate GDP (at PPP) as an indicator of the relative size or strength of whole economies—as a factor, for instance, in assessing geopolitical power. China’s GDP at PPP recently passed that of the United States. But such calculations neglect the fact that PPPs take account of the systematic tendency of poorer countries to have lower prices. If this tendency is attributable to the lower productivity of lower-income economies in the production of internationally traded goods and services (as implied by the most
widely cited explanation, attributed to Balassa [1964] and Samuelson [1964]), the PPP adjustment goes too far, attributing greater economic strength to these countries than it should. Removing this systematic part of the PPP adjustment reorders measures of economic size, placing China behind both the United States and the European Union (even without the United Kingdom), for example.

Furthermore, PPP adjustment buys less than might be supposed. It is often just the first stage in putting data on a comparable cross-country basis. The well-known shortcomings of GDP affect countries to differing degrees. If cross-country comparisons are to be robust, they need to take account of factors such as environmental degradation and the globalization of production. In at least 10 countries, the required resource depletion adjustment for environmental degradation amounts to more than 15 percent of GDP. The far-fetched 25 percent GDP growth rate in Ireland in 2015 reflects the role of multinational corporations.

THE NEW DATA

The ICP began its price collection exercise half a century ago, in order to facilitate international comparisons of economic quantities measured in national monetary terms free from the distortions that can result from exchange rate volatility. The heart of this exercise involves the calculation of artificial exchange rates (PPPs) that would, if they were in effect, equalize the average prices prevailing in every country.¹ Using PPPs instead of market exchange rates allows the real value of GDP and related income and production data to be compared on a more robust basis. PPPs are an essential part of the toolkit of international economic analysis.

Every few years, a huge worldwide data collection exercise is conducted in order to compute PPPs and PPP-adjusted macroeconomic statistics for virtually all countries in the world. Between updates (which until now had been conducted at six-year intervals but will henceforth be conducted every three years), annual data are extrapolated in a simple way.

The results of the latest benchmark round (relating to 2017) were published in May 2020 (World Bank 2020). They produced fewer surprises than the results published in 2008 (for the 2005 benchmark) and 2014 (for 2011) and have been greeted with much less controversy. The average change in the estimate of GDP (at PPP) was about ½ percent. The estimated extreme poverty headcount is also revised by just a fraction of 1 percent, according to Atamanov et al. (2020). Still, GDP estimates for about 20 economies changed by at least 10 percent.

Country coverage in the early years of the ICP was limited; it increased from 10 countries in 1970 to 60 by the turn of the century (Deaton and Aten 2017, Johnson et al. 2013). Before the 2005 ICP, data collection operated on a shoestring, resulting in operational difficulties and delays. Revisions made between the benchmark data collection exercises were sufficiently large to overturn the results of several influential econometric studies carried out on prerevision datasets (Johnson et al. 2013).

¹ It is necessary to distinguish between the synthetic exchange rates known as PPPs and the theory of purchasing power parity. The PPP theory, in essence, hypothesizes that market exchange rates tend to converge to these PPPs.
Among a growing number of entities involved in the program, the World Bank took on an increasingly large role in what had become a very demanding international data collection effort. The benchmark price collection exercise for 2005 covered 146 countries (including, for the first time, China). When published, in 2008, the new price data—and changes in the methodology for constructing the PPPs from the raw data—resulted in a substantial downward revision in estimated income and poverty levels for almost all developing countries. The two largest developing countries, China and India, were reestimated to have per capita incomes that were about 40 percent and 35 percent, respectively, below the earlier estimates. Scholars debated the results, with some doubting whether the low estimate for China and India could be consistent with reported GDP growth rates for those countries over the previous decades. Projecting the new estimates of income backward to 1950 would imply that millions of Indians were then at income levels below subsistence (Deaton and Heston 2010, Ravallion 2018).

Six years later, the 2011 benchmark partly reversed the surprises, both for China and India and more generally. Estimated price levels in lower-income countries were much lower than previously thought. Accordingly, average real incomes in those countries were much higher. The number of people below the World Bank’s extreme poverty line (which is set in PPPs) declined by about half, once more threatening to “play havoc with the poverty estimates” (Deaton 2001). The World Bank promptly raised its extreme poverty line from $1.25 to $1.90 a day (to be converted into national currencies at PPPs). In light of the volatility of the estimates for successive benchmark datasets, Atkinson (2016) suggested that, for policy purposes, the international poverty line as measured in national currencies should not be revised in the light of new rounds of the ICP; however, this has not been done.

WHY ARE PPPs IMPORTANT?

To illustrate how important it is to make these adjustments, consider the case of Argentina in recent years. Converted at published exchange rates, Argentina’s per capita GDP was about $14,600 in 2017, or about 24 percent of the corresponding figure for the United States. Two years later, its GDP had fallen to $10,000, or 16 percent of the US level.

If per capita GDP is considered a measure of economic performance, should Argentina be thought of as closer to 24 percent of the US level or 16 percent?\(^2\) Was there really that deterioration in just two years? Such questions led to the construction of PPPs as a substitute for market exchange rates for making comparisons.\(^3\)

Comparing US GDP in dollars with, say, Argentine GDP in pesos requires the use of an exchange rate, given that national accounts statistics are constructed on the basis of domestic prices. Over the long run, market exchange rate movements tend to be aligned with relative price inflation. Sudden exchange rate movements can leave domestic prices, especially the prices of nontraded goods

\(^2\) The peso declined further during 2020.

\(^3\) The World Bank’s extreme income poverty threshold is a component of seven of the United Nations Sustainable Development Goals. It is one of the three dimensions of the United Nations Development Programme’s Human Development Index.
and services, well out of line with those abroad, however. Furthermore, when exchange rates are administered or pegged, they can remain far from their long-term equilibrium relation with relative prices for years, especially if exchange controls are in effect. Comparing GDP across countries at market exchange rates can therefore be misleading.

In Argentina, for example, most of the apparent decline in its GDP between 2017 and 2019 can be traced to the plummeting of the market value of the Argentine peso in those two years. Although there was also rapid inflation in Argentina during this period (prices more than doubled), prices had not caught up with the decline in the market value of the peso. The latest PPP-adjusted data, for 2019, put Argentina’s GDP at 35 percent of the United States’, down from 39 percent in 2017—a much more modest decline.

Several other examples of crises over the past two decades illustrate why PPPs are important.

Figure 1 shows per capita GDP at PPP and at market exchange rates for six volatile economies, each shown as a percentage of the US figure. The PPP series is much smoother in terms of percentage changes. The decline in market exchange rates greatly overstates the severity of the economic collapses experienced by each of these countries (severe though they were).

A stark example of the divergence in estimates based on the two measures is Argentina (figure 1, panel a), where the decline in per capita real income relative to the United States in 2001–02 was about 12 percent based on PPP and 65 percent based on market exchange rates.

The crisis in Indonesia in 1998 reduced per capita real income by about 17 percent relative to the United States—not the close to 60 percent indicated by market exchange rates. A similar story applies to the Russian economy in 1998–99. The volatility of the Turkish series (measured by the coefficient of variation) is about 32 percent in the market exchange rate series but only 20 percent in PPP. In 2008, after its massive two-decade decline, Zimbabwe’s income level was nevertheless almost twice as large (in PPP terms) as implied by the decline in its GDP at market exchange rates. In contrast, Venezuela’s per capita income relative to the United States grew by about a third between 2003 and 2010—not by the 240 percent given by market exchange rates.

Correcting large exchange rate fluctuations that do not reflect changes in relative prices is important if serious errors are not to be made in making international financial and economic comparisons. PPPs are an essential ingredient in all such comparisons.

But, apart from removing distortions caused by market exchange rate movements, PPPs also remove the impact of other factors affecting relative prices, especially systematic productivity differences. For some purposes—especially when one is looking at the overall economic size or strength of national economies—this may be going further than is desirable.

**INTERPRETING THE ECONOMIC SIZE OF LOWER-INCOME COUNTRIES**

According to the 2020 ICP report announcing the new series *Purchasing Power Parities and the Size of World Economies*, China’s GDP in 2017 was marginally larger than that of the United States, with India in third place, all measured at the new PPPs. (At market exchange rates, the rankings would be quite different,
with the United States ahead of China, and India coming in sixth place, after Japan, Germany, and the United Kingdom.) But the use of PPPs in this context removes not only the distortions resulting from mispriced exchange rates but also price differentials that are attributable to other factors. One of these factors, productivity differentials, arguably should not be removed if one wishes to compare the overall economic size or strength of a country or region.

Source: World Bank, World Development Indicators.
Lower-income countries tend to have low average wages and relatively low prices for nontraded goods and services relative to high-income countries. This systematic effect is sizable and has been evident ever since data were first collected a half-century ago (figure 2). The most commonly cited explanation for this pattern, known as the Balassa-Samuelson effect, relates to supply conditions. According to it, relative wage levels across countries will be systematically related to productivity levels. Where average labor productivity is low, wages will also be lower. International competition is likely to equalize the prices of traded goods, but lower wage rates in a lower-income country will be reflected in lower prices for nontraded goods and services. If, however, as the Balassa-Samuelson theory asserts, low average productivity (in tradables) is the source of a systematically lower price level in lower-income countries, recalculating GDP

4 The pattern is not strong at low income levels; it becomes more evident as income rises (Hassan 2016). Three other factors strongly influence the cross-country pattern of price levels as reflected in the ratio between PPPs and market exchange rates: whether the country is a petroleum exporter, whether the country is rated as having a high level of corruption, and whether a country is geographically remote. Each of these factors tends to increase the price level in developing countries (Cheung, Chinn, and Nong 2017).

5 Demand conditions could also be relevant. The relative price of services will be driven up, for example, if consumers’ preference for nontraded services increase with income (De Gregorio, Giovannini, and Wolf 1994). Yet another cause could be the relative supply of labor, which is high in most developing countries, driving down the price of labor-intensive services (Bhagwati 1984). Econometric research does not unambiguously settle which of these factors dominates (Devereux 2014).
at an artificial exchange rate that removes this effect may result in exaggerating their economic size, productive capacity, and importance in international economic relations.

This result suggests that it would be advisable to move only part way from market exchange rates to PPPs, by removing that part of the adjustment that is thought to be attributable to productivity differences. A simple and very rough way of approximating this share is to subtract the estimated effect of per capita income on PPP, as done in the literature. Assuming that it is indeed productivity that is at work, bias in the measure of an economy’s productive capacity could be avoided by adding back that factor to each country’s PPP estimate to get a simple productivity-adjusted PPP. This adjustment does not directly address each country’s productivity, only the average productivity gain as income per capita increases across the world. The resulting adjusted data, which seem better suited for comparing the economic strength of different countries, significantly reorder the ranking of even populous countries and economic areas.

The results of such an adjustment for average productivity in tradables are striking (figure 3). The United States retains its top position, and the European Union (even without the United Kingdom) is just ahead of China. India is only just ahead of Japan in fourth position, instead of being 60 percent larger than Japan. In sixth place is Russia, well ahead of the aggregate for Sub-Saharan Africa.

6 The adjustment amounts to swivelling the data in figure 2 until the straight regression line is horizontal. The slope of the regression on the 2017 data is 0.20. Based on earlier benchmark data, Hassan (2016) and Cheung, Chinn, and Nong (2017) find a steeper slope of 0.24. They also propose (and prefer) a nonlinear version, created by adding a quadratic term to improve the fit for lower-income countries (as illustrated in figure 2). Using their measure would further accentuate the changes for the largest national and regional economies.
(which is larger than Russia in the unadjusted PPP series). This crude adjustment lacks the subtlety of the ICP exercise, but it is a convenient way of highlighting a major potential pitfall in using GDP at PPP to compare the economic size or strength of different countries.

**METHODOLOGICAL ISSUES**

Controversy has persisted about the best way of calculating an index of cross-country price differences. Comparing the welfare of people with different preferences facing different prices is “close to being impossible in theory,” according to Deaton and Aten (2017).

One pragmatic approach has been to take an average of all bilateral relative prices. An alternative approach is to make the calculation relative to a single set of reference prices and quantities—a single artificial reference or “average” country, as it were. The choice between these approaches—and additional methodological choices within each approach—can strongly affect the estimated PPPs (Neary 2004, Oulton 2012). To the extent that the averaging process moves the PPPs closer to reflecting expenditure weights of high-income countries (as appears to be the case), the approach adopted by the ICP tends to understate the gap between high- and lower-income countries.

Collection of basic price data has also been a challenge. Though greatly improved over the years, the process remains imperfect. Some goods that are common in advanced economies are not available in some developing countries or available only in a handful of outlets catering only to the rich or to foreign visitors. Should these prices be included, or should an alternative approach be adopted?

A highly complex system has been adopted to deal with this and similar problems. But controversy remains. For example, there may be a bias toward internationally traded goods in selecting items to be priced for the comparisons, as Ravallion (2018) has suggested as a possible contributory factor to overstating price levels in lower-income countries.

Some observers, including Reddy and Pogge (2010), have criticized the use of PPPs based on an average household consumption basket (rather than a basket reflecting consumption by the poor) for defining poverty lines. But a recent study of 16 African countries finds that in most of them, the prices paid by the poor are lower than average (Dabalen, Gaddis, and Nguyen 2020). However, Argente and Lee (2020) show that during the Great Recession in the United States, inflation for the poor was higher than for others. The use of market exchange rates would not help here: PPP adjustment is far from the last word on defining poverty lines, but the factors that it takes into account are important. Poverty is a multidimensional concept that is not fully captured by measures of the purchasing power of cash income (Atkinson 2019).

---

7 Germany, the United Kingdom, France, and Italy all fall in the top nine countries in the new ranking. Brazil slips to 10th place, instead of jockeying with the United Kingdom and France for 7th place in the unadjusted PPP ranking.

8 This point is different from the point that PPP-adjusted data are not best suited to measure national real GDP growth rates, especially between benchmark years, when interpolation is based on national data (Johnson et al. 2013).
PITFALLS IN COMPARING AVERAGE INCOME LEVELS

The productivity issue does not arise when the ICP data are used for comparing per capita income across countries. Yet there are pitfalls here, too. GDP has shortcomings as a measure of welfare (Coyle 2014). The impact of these shortcomings is highly differentiated across economies, likely contaminating cross-country comparisons more often than comparisons for a single country over time.

Consider, for example, environmental degradation. In 10 countries, the World Bank’s estimates of natural resource depletion exceed 15 percent of GDP; in some 50 countries, it exceeds 5 percent. Many of these countries are small producers of primary goods, but they include the Democratic Republic of the Congo, with its 90 million people, as well as more prosperous countries such as Chile, Kuwait, Peru, Qatar, Russia, and Saudi Arabia. In contrast, in about 50 countries such depletion is estimated at less than 0.2 percent of GDP. Neglecting such large differences can clearly skew comparisons.

The ambiguous geographical location of the activity of some of the largest multinational corporations further distorts international comparisons of GDP. This problem is evident in the data for Ireland in recent years.

Ireland is home to many multinational corporations, especially in the fields of information and communication, pharmaceuticals, and aircraft leasing. Depreciation of the assets (physical or intangible) owned by the Irish units of these firms enters into Ireland’s GDP. In 2015, a jump in these assets resulting mainly from a transfer of ownership resulted in a 25 percent increase in Ireland’s real GDP. Although the recording of these activities is consistent with international statistical standards in the UN System of National Accounts, such movements wreak havoc on most conventional uses of GDP. As a result, GDP per capita at PPP is no longer a meaningful measure of living standards for Ireland.

Fortunately, the ICP produces specific PPPs for another variable—actual individual consumption (AIC), one of the major subaggregates of GDP—which is closer to a consumption-based measure of economic living standards and which is free of the distortions caused by the activities of multinational corporations. AIC includes both household consumption and that part of government consumption (and consumption by nonprofit institutions) that is provided directly to individual households. The ICP regards AIC as a “measure of average material well-being.” Using AIC per capita in place of GDP per capita (and ignoring economies with populations of less than 1 million) moves Ireland from fourth place in the world (at 130 percent of the US level, ahead of Switzerland and Norway) to 21st (at 60 percent of the US level), a much more

---

9 The agreed upon methodology for integrating resource depletion into the national accounts is known as the United Nations System of Environmental-Economic Accounting (SEEAA). Such accounts have not yet been prepared on a PPP basis (Lange, Wodon, and Carey 2018).

10 The ICP publishes PPPs for 43 subaggregates of GDP. Some, such as gross capital formation and government consumption, are large. Others, such as construction and food, are moderately large. Some subaggregates, such as meat or fruit, are very specific (World Bank 2020).
realistic ranking. It would be good if headline writers and Twitter users would use AIC rather than GDP for countries in which activities of large corporations disengaged from the welfare of households are especially high.

**CONCLUSION**

Having a good statistical basis for cross-country comparisons is essential for good policy analysis. Appropriate interpretation and use of the data are equally important.

The scale of the conceptual and practical problems involved in constructing PPPs for all of the economies in the world should not be underestimated. But acceptable practical solutions have been devised, and even if the methodological debates have not been fully resolved, they have been quieter of late. The PPP data effectively neutralize the distortions caused by exchange rate volatility. Other important structural differences affecting the interpretation of GDP comparisons across economies, including environmental degradation and the statistical impact of multinational corporations’ location decisions, should not be neglected.

The widely accepted Balassa-Samuelson theory, which interprets the lower prices that prevail in lower-income countries as attributable to low productivity in tradables, has an important neglected implication for cross-country comparisons of national economic size and strength. Taking it into account significantly reorders the ranking of economies.

As the world economy restructures after the COVID-19 pandemic, patterns of consumption and production are likely to change significantly. With sharply changing patterns of expenditure in 2020 and likely beyond, national and international price indexes may need more frequent revision (Diewert and Fox 2020).

---

11 Other countries that slide down the ranking when one moves from GDP to AIC include Qatar, Singapore, and the United Arab Emirates. The appendix provides detail on the Irish case.
APPENDIX

STATISTICAL DISTORTIONS IN THE IRISH NATIONAL ACCOUNTS CAUSED BY ACTIVITIES OF MULTINATIONAL CORPORATIONS

Ireland is a prosperous country—but not as prosperous as is often thought, because of the inappropriate use of misleading, albeit conventional, statistics. PPP adjustment takes account of the high prices prevailing in Ireland, but other distorting factors are even more important.

For years the large profits reported by foreign multinationals have affected Irish GDP growth rates. Until relatively recently, most of these profits were attributable to foreign parent companies. They did not distort another standard international measure, gross national income (GNI), which excludes profits going to foreign owners.

Two other large and even more volatile distortions induced by activities of multinationals have grown substantially in recent years. They affect both GNI and GDP. The larger of these distortions relates to the very large capital assets owned by multinationals operating in Ireland. These assets include intangible intellectual property assets (such as patents) of information technology companies and the large aircraft fleets of resident leasing companies. The acquisition of these assets from abroad does not contribute to Ireland’s GDP. Once acquired, however, the depreciation of the assets is included in both GDP and GNI, because depreciation is exactly what the G (for gross) means. It was the jump in depreciation during 2015, following the decisions by multinational companies to move a large block of intellectual property into the ownership of resident affiliates, that led to the 25 percent GDP growth rate in Ireland that year.

The other GNI distortion relates to multinationals that have “redomiciled” (i.e., moved their global headquarters to Ireland despite having little domestic economic activity in Ireland). Their undistributed profits—which exceeded 5 percent of GNI in 2012 before falling back somewhat—are also included in GNI, even though their shareholders are mostly nonresident (FitzGerald 2018).

To help analyze the Irish macroeconomy, the Irish Central Statistics Office (2016) developed a measure known as GNI*, which excludes these distorting factors.12 In 2019, GNI* was about 40 percent below GDP (in most countries, GDP and GNI are about equal). Applying this percentage reduction to GDP to make a rough international comparison (using PPP-adjusted data) reveals that Ireland was not in second place in the European Union (after Luxembourg) but instead in eighth place, a ranking similar to that obtained by using AIC per capita at PPP.13

---

12 GNI* subtracts from GNI factor income (mainly profits) of redomiciled companies and depreciation on aircraft leasing and imported intellectual property (R&D service imports and trade in intellectual property).

13 In the 2019 Human Development Index (HDI), Ireland ranked third in the world, behind Norway and Switzerland. But the ranking was affected by the distortion caused by multinationals, as the HDI uses the ICP numbers for GNI, along with health and education indicators. Scaling down GNI by the ratio of GNI* to GNI (0.772 in 2019) reduces Ireland’s HDI score and moves Ireland down from third place to ninth in the global ranking.
REFERENCES


This publication has been subjected to a prepublication peer review intended to ensure analytical quality. The views expressed are those of the author. This publication is part of the overall program of the Peterson Institute for International Economics, as endorsed by its Board of Directors, but it does not necessarily reflect the views of individual members of the Board or of the Institute’s staff or management.

The Peterson Institute for International Economics is a private nonpartisan, nonprofit institution for rigorous, intellectually open, and indepth study and discussion of international economic policy. Its purpose is to identify and analyze important issues to make globalization beneficial and sustainable for the people of the United States and the world, and then to develop and communicate practical new approaches for dealing with them. Its work is funded by a highly diverse group of philanthropic foundations, private corporations, and interested individuals, as well as income on its capital fund. About 35 percent of the Institute’s resources in its latest fiscal year were provided by contributors from outside the United States.

A list of all financial supporters is posted at https://piie.com/sites/default/files/supporters.pdf.