Moatsos, Michail

Working Paper
Why PPP exchange rates should be avoided in global poverty estimates

Suggested Citation: Moatsos, Michail (2020) : Why PPP exchange rates should be avoided in global poverty estimates, ZBW – Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:
http://hdl.handle.net/10419/218972
Why PPP exchange rates should be avoided in global poverty estimates

Michail Moatsos*

April 15, 2020

Abstract

Purchasing Power Parity (PPP) exchange rates work appreciatively for comparing economies across the globe, instead of the standard market exchange rates. PPPs come closer to represent the relative size of the economies because they correct for non-tradeables that are relatively cheaper in less developed countries. However, those rates are constructed for comparing countries, or their households, in sum. Thereby when they are used to compare sub-groups, e.g. the poor in each country, their methodological foundations are stretched beyond their specifications. This paper highlights the often neglected issues that are raised from this standard practice in global poverty measurement.

*Author affiliation: Department of History and Art History, Utrecht University, Drift 6, 3512 BS, Utrecht, Netherlands
1 An overview of the dollarized poverty line issues.

“Poverty estimates for a country should not change simply because other countries’ consumption patterns or price levels have changed, nor because the consumption pattern or price level of goods that are not needed to avoid poverty have changed. A method of measurement that fails to satisfy this requirement is flawed.” Reddy and Pogge (2010)

The estimation of PPP exchange rates is a data intense, and methodologically demanding exercise. For this purpose a worldwide collaboration among the World Bank and national statistical authorities sets-up the framework of the International Comparison Program (ICP). For the latest round in 2011, 199 countries are covered, 53 more than the previous round in 2005. This entails an enormous methodological and statistical endeavor, that includes considerable improvements in coverage and homogenization of the various processes compared to previous rounds. The question that remains though is not related so much with how well is the design of an ICP round, and the extend of the resources allocated to it, but if the resulting PPP exchange rates, and the methodology that underlines them, are appropriate for use in global poverty measurement.

The World Bank warns about the application of PPP rates for poverty estimates, by acknowledging that PPP estimates “may not reflect the expenditure patterns of the poor” (TheWorldBank, 2007). The reason for issuing this warning is that the PPP estimates are calculated either for the economy as a whole or for all the households, and therefore in neither case reflecting the expenditure patterns of those living in conditions of poverty. A related example pointed out by Deaton (2001) is the 0.1 percent poverty rate in Thailand in 1997, which World Bank’s Chief Economist at that time, Joseph Stiglitz, has cited as “one of the consequences of the Asian economic miracle”. Deaton argues that this finding is a result of unsuitable PPP conversions, instead of an economic miracle. The point being that the unsuitability of PPP exchange rates is casting reasonable doubts over poverty estimates that make use of them. Similar warnings, albeit in different format, have been issued for the latest round of PPP estimates using 2011 as the benchmark year, that was published in June of 2014 by the then chief economist Kaushik Basu. According to the Brookings Institution, preliminary calculations using the 2011 PPPs brings an immediate poverty reduction of between 25-to-50+ percent, depending on adjusting the poverty line or not; their method however ignores some fundamental steps thus producing marked differences.

---

1Emphasis in the original.
2The next ICP round is conducted in 2017 but the results are not available yet at the time of writing this thesis, and are expected in late April 2020 (ICP Highlights, Issue 44).
3World Bank’s Understanding PPPs.
4What Do New Price Data Mean for the Goal of Ending Extreme Poverty? at Brookings
Overall, compiling a consistent PPP dataset is a challenge in itself, but taking the next step in building PPP datasets that “reflect the relative price levels of the goods and services faced by poor consumers” (Aten and Menezes 2002) is an additional challenge which some scholars suggest that won’t be a feasible option. Klasen et al. (2016) conclude that “it would be best to consider alternatives to the current reliance on ICP rounds and the resulting PPPs.” Without going into the underlying formulas, I demonstrate in the following subsections a number of core reasons why one should indeed—as Klasen et al. recommend among others—abandon the use of PPP exchange rates in poverty estimates.

1.1 Consumption patterns

In their seminal article Ravallion et al. (1991, p.5) state that “[i]deally one would like to construct new PPP rates for the prices most relevant to the absolute poor, in which the prices of food-staples would clearly carry a high weight”. Before them Ahluwalia et al. (1979, p.305) already acknowledged that the application of the Kravis ratio (which is how PPPs were called at the time) is more appropriate than market exchange rates in global poverty research, but other problems arise that replace the problems addressed. One concern relates to the likelihood that PPPs vary among various income groups within a country. Another element that concerned them was that the switch from market rates to PPP rates is based, among other things, on the undervaluation of services in developing countries, in turn this may well mean that “official exchange rates understate incomes of the rich more than of the poor”, since services are consumed more by the higher income groups within those countries. Averaging out this into a single PPP rate simply turns a blind eye to the problem, as it assigns to the poor part of the additional (compared to market exchange rates) purchasing power that should be attributed to the richer.

Theoretically, PPP comparisons may have a potential to become ideal once all the products in the calculations are representative for all participating countries. Obviously, this holds when one wants to compare countries. When the goal is to investigate how specific population groups in each country compare with their corresponding groups in all other countries, then again calculations should include representative consumption elements of those groups. In any case, the practice of the World Bank in estimating consumption PPPs ignores this point. Pogge (2013) provides an illuminating numeric example. Imagine a simplistic world where there are only three commodities: necessities, discretionary, and services. Assume that the prices for these commodities in country X are 5, 6 and 1 respectively, and in numeraire country A respective prices are 3, 4 and 9. Pogge calculates the PPP to be 1.55, meaning that each local currency unit (LCU) in country X is equivalent, in PPP terms, to 1.55 numeraire LCUs. If one however only focuses in necessities consumption, as one would
when identifying the poor, X country’s LCU worths 0.6 numeraire LCUs. The implication of this, as they put it, is straightforward: “The Bank’s reliance on general consumption PPPs ensures that, wherever the actual price of necessities is higher than what such PPPs suggest, many who are very poor, relative to what they really need to buy, do not show up in the Bank’s extreme poverty statistics” (ibid).

As said, the above example is very simplistic, and assumes further that the three commodities consumed in the two countries are identical, and representative. More often than not, neither of the two holds exactly. Even when comparing similarly poor countries, the products that are necessary for survival may well be country specific. Deaton’s (e.g. in 2010 and 2013) favorite related example is that of teff in Ethiopia, which is rarely used anywhere else, and tofu in Indonesia. Both are basic food stuff consumed by the poor in those countries. But when one wants to compare the poor in the two countries, pricing appropriately those products is simply impossible, as there is no teff in Indonesia and no tofu in Ethiopia. There are methods to estimate a “reasonable” price by regression, but these estimates cannot correct for the fact that any estimated price does not represent anything real. Those prices are simply statistically convenient structures that make the estimation of PPP exchange rates possible. The bias can work either way in those estimates.

Going back to the teff and tofu example, let’s further assume that both of these products are consumed in country X, and then follow the steps taken by Deaton and Heston (2010): the price of teff in Ethiopia is \( P_{\text{Ethiopia}}^{\text{teff}} \), the price of tofu in Indonesia is \( P_{\text{Indonesia}}^{\text{tofu}} \), and in country X teff is priced at \( P_{\text{X}}^{\text{teff}} \) and tofu at \( P_{\text{X}}^{\text{tofu}} \). Then the imputation method would give a parity of tofu relative to teff as the product of \( P_{\text{tofu}}^{\text{X}} / P_{\text{tofu}}^{\text{Indonesia}} \) (the parity of tofu between country X and Indonesia) and \( P_{\text{X}}^{\text{teff}} / P_{\text{Ethiopia}}^{\text{teff}} \) (the parity of teff between Ethiopia and country X). The conclusion of Deaton and Heston is that this estimation “is certainly arbitrary in the sense that the parity between two countries depends entirely on information from third countries”. This is of course a problem related to the nonexistence of important products in some third countries. The problem of course persists even in milder versions related to products that are less than representative of consumption patterns in a country. Again biases are introduced in the estimates. Those problems tend to be augmented when one needs to compare countries with dissimilar patterns of relative prices and expenditures.

Besides tentative differences in staple food consumption patterns, other GDP components that are “comparison-resistant” include government provided services, health care, education, construction, and house rental. According to Deaton and Heston, due to the importance of those GDP components their treatment can affect country wide and region wide PPPs. For the house rental component, the cases of Asia and Africa are treated differently than other regions. The implication of this difference in treatment is that housing volumes cannot be meaningfully compared between countries within and outside of those regions. For Ghana, Chad, and Malawi they estimate that the divergence in
PPP rates when including or excluding the rental category can be close to 10 percent. Deaton (2010, p.14) estimates that using a price-wise neutral treatment of the rental component a reduction of poverty count for 2005 “by more than 100 million people”.

When calculating PPPs the more one commodity is consumed, the more its price will influence the final PPPs. Pogge (2013) maintains that PPPs are influenced too much by the prices of commodities that are irrelevant to absolute poverty avoidance, such as luxury goods and services. Inversely, PPPs are influenced relatively little by the relevant necessities to those who live in conditions of poverty. Along the same lines Aten and Heston (2010) conclude that available household consumption PPPs is an improvement compared to the GDP (or economy wide) PPPs, as they exclude investments and government expenditure. However, still the average consumption patterns differ with the patterns of those that struggle for survival in conditions of extreme poverty. They suggest that one could focus on consumption patterns of the poor, and the respective prices they face, but the problem is hard to solve due to unavailability of such data. Ravallion (2010) criticizes the idea of pricing a single basket of goods implied by Reddy and Pogge (2010) and Aten and Heston (2010) on the grounds that consumption patterns differ among the poor in different countries (like in the tofu and teff example). However, this criticism holds only in the extreme case of rigidly fixed recipe in the space of products for pricing the consumption basket that would underlie a cost of basic needs poverty line, and as Allen (2013) argues, there is no reason why the baskets cannot adapt, e.g. to local price specifications by choosing the cheapest combination of products to achieve a specific goal.

1.2 Country (ir)relevance

PPP estimates are more reliable and accurate when the participating countries have similar consumption patterns and similar economic structure. The more the countries differ in these perspectives the larger the resulting concern for the PPP estimates.

Of particular concern are the unrepresentative high-end prices in poor countries, when one constructs PPP rates that include a mix of poor and rich countries. When one is estimating global absolute poverty figures using PPPs it follows that –as discussed above– the number of poor in one country will fluctuate based on the change in prices in a third country, even if nothing changed in the investigated country and the numeraire country (Reddy and Pogge 2010). On this topic Deaton and Heston (2010) uses the example of consumption of wine. Considering the case that the expenditure share for wine in Cameroon is small, it is the case due to the application of GEKS calculations (see relative subsection below) that the price of wine in Cameroon will attract some of the price from France or other rich countries that consume more expensive wine.

The approach of linear programming is used throughout this thesis to account for this issue. See the following empirical chapters on poverty for details.
This would imply an overstatement of prices in Cameroon, and an understatement of its real GDP in PPP terms. The effect will not be that great since the consumption of wine in Cameroon has a small share. For other products with larger share the impact in understatement of real GDP would be larger.

A related and rather unexpected issue is that of the global political balance in getting the final ICP calculations. As Deaton and Heston (2010, p. 18) discusses the participation of Eurostat in the ICP rounds since 1980 is made conditional on ICP respecting the regional PPPs as estimated by Eurostat. This calls for additional fixity concerns that are political and not statistical in nature. Deaton estimates that without imposing this type of fixity constrains one gets a 6.6 percent higher real GDP for China for the 2005 PPP round.

1.3 The main point of Reddy and Pogge criticism on the use of PPPs

The final step in the calculation of PPPs involves an adjustment that fulfills the so called “transitivity” requirement among the PPP rates for the various countries. This means that the PPP rate between say India and USA, for example, should be the same if it is estimated directly or via a group of third countries, say via Luxembourg. This final step influences the PPP rates not only with respect to the relative prices and spending structure of the numeraire country, but also with those from every other country (Pogge, 2013). Or in the words of Pogge (2013) “The fact that an income suffices to meet basic human needs [in one country] is no assurance, then, that a PPP equivalent income in another country is similarly sufficient. In poor countries, prices of necessities are often higher, and prices of services lower, than what the PPP to the US dollar would suggest”.

This does not mean however that without imposing transitivity, the bilateral PPP rates would be more useful, as still commodities and consumption patterns of the numeraire country would influence the poverty status in all other participating countries. This relativity in the World Bank method cannot be accounted for. The dollar-a-day methodological approach supports the idea that it is possible to pinpoint a single common poverty line in a common denomination, when at the same time the PPP process itself, and the problems inert to the PPP estimation method, make evident that such equivalence is biased (towards an unclear direction), and most likely the bias is different for each country. This argument holds for every PPP dataset at its benchmark years. If one moves the comparison beyond the benchmark year an additional bias via the application

---

6Of course in this example the issue of quality is not treated, but quality has similar problems as in the case of consumption patterns discussed here (Deaton and Heston, 2010).

7This translates to the methodological concern quoted in the beginning of this section, whether the World Bank categorizes a person as poor or not according to the iPL, it is not only affected by the available means of the person and the prices that person faces, but also on the prices and consumption habits of all countries participating in the ICP round. Deaton (2001) offers a relief with respect to this point, by indicating that the PPP rates pre- and post the imposition of the transitivity constraints are very similar.
of domestic CPIs, that have different consumption structure than the one implied by the ICP, and the prices that the poor face, affect the bias further. The further we go from the benchmark year, the PPP estimates become even less reliable for global poverty research.

The assiduous effort of Deaton and Dupriez (2009) in calculating PPP rates relevant for the poor—although has to some extend attenuated the problems mentioned above—for a number of reasons the PPP for the poor (P4s) they provide do not address the issues fully. On the one hand, as the authors recognize, there are problems that relate to the availability of data below their "basic headings". This means that the data they used in their calculations come in some form of aggregates and do not include the actual micro data of commodity prices and volume from the 2005 ICP round. On the other hand, consumption patterns and country irrelevance problems arise even if one focuses only for the developing world, excluding e.g. OECD countries, or considering only those basic headings that are arguably more relevant to the less poor. The fact that on aggregate PPPs and P4s give very close results according to Chen and Ravallion (2010), does not mean that the differences in a per region or per country basis are negligible as found in Deaton and Dupriez (2009, Table 16). The coincidence on aggregate brings no guarantees that would be so in forthcoming ICP rounds, especially if the underlying data below the basic headings become available for independent evaluation.

1.4 PPP estimates using GEKS & GK

There are two main sources of PPP estimates, the Penn World Tables and the World Development Indicators. PWT was GK-based (Geary-Khamis) and WDI is GEKS-based (Gini-Eltető-Kőves-Szulc). In a sensitivity analysis Ackland et al. (2013), have shown the impact of different methods in calculating the PPP rates. They conclude that the GK method understates the number of the global poor relative to the GEKS method. This is mainly due to the higher purchasing power attributed to the relatively poor countries. From

---

8To date there is no similar attempt to that of Deaton and Dupriez (2009) for the 2011 ICP round.
9As of version 8.0 they changed their approach to a combination of methods. For the latest method underlying PWT see Feenstra et al. (2015).
10I follow the GEKS convention introduced by Deaton, in recognizing the priority of Corrado Gini in the conception of the method.
11This is achieved by the way the international price vector is computed, that brings the vector closer to the prices prevailing in the rich OECD countries, rather than the less affluent ones. This approach ignores the substitution for cheaper commodities that takes place in the poorer countries. Thus the income in poor countries is overstated, an effect dubbed as the Gershenkron effect. Further, in the GEKS method, the GEKS quantity index is the geometric mean of Fisher quantity indexes. In turn the Fisher quantity index is the geometric mean of the Laspeyres and Paasche quantity indexes. It is known that Laspeyres index shows a propensity to overstate the income of rich countries relative to the poor, and the opposite is true for the Paasche index. Apparently there is no guarantee that the Fisher index will be bias free as a result of taking their geometric mean. However, it is the case that the bias is smaller than the one introduced by the GK-method, as also Ackland et al. (2013) demonstrate.
PWT version 8.0, some steps to addresses those concerns over previous versions have been made. Reddy and Pogge (2010) argue that both methods may introduce artificial declines in poverty rates. For the GEKS method this bias operates via the rising share of services in consumption, because those services are relatively cheaper in poor countries compared to the rich. This brings about a decline in PPPs, and thus an artificial decline in poverty as a mechanical product of time. The bias in the GK method is driven by shifts in consumption from tradeables (e.g. food items, manufactures, etc.) to non-tradeables (e.g. housing, local services, etc.) that result in reducing the PPP of poor countries (and as a result also reduce their poverty rates as measured by the dollar-a-day method).

1.5 CPI implications

The standard practice in poverty estimates is to apply the PPP exchange rates for the PPP benchmark year for each country. Then in order to get poverty estimates for any previous, or following, year the domestic CPI is applied for each country. At the same time it is widely accepted that the purchasing power equivalence does not necessarily hold with this CPI application (Chen and Ravallion, 2010). As discussed in Deaton and Heston (2010, p.27) the application of CPI will not match the PPP exchange rates derived in the next (or any previous) benchmark year. This issue obtains, among other reasons, because of: (i) the differences between items priced for the domestic index and the items priced for the ICP rounds, (ii) the differences in the geographical coverage where price collection takes place in domestic and ICP rounds, and (iii) because this step takes place for each country separately and no transitivity constraint is applied as in the calculation of PPP exchange rates.

The implication of this can also be understood from the perspective of the iPL. To come up with the poverty rate for each country, the iPL is converted to local currency units (LCUs) using PPP rates for the benchmark year. To estimate the LCUs that correspond to the iPL line for the year before the benchmark its value is discounted by the domestic CPI rate, and so forth for any previous year. This is done for all countries under inspection. In principle then the iPL moves into different trajectories for each country separately for the reasons mentioned above. The larger the distance from the benchmark year the wider the implications of domestic CPI application on the global poverty estimates. Clearly, this has a stronger impact in more long run estimations of poverty, and the further away we move from the ICP benchmark year, the larger becomes the concern regarding the assumption that an iPL maintains purchasing equivalence all over the world. Ideally for the dollar-a-day method, in this regard, one should have a new ICP round every year producing PPP exchange rates that would be used only for that year.

These are the reasons why Chen and Ravallion (2001) argue when comparing the iPLs between 1985 and 1993 rounds, that "\footnote{Therefore, the iPL in non-benchmark years is not the international one any more, as it has been domesticated by the CPI application and conversion to LCUs.} in effect, the whole structure...
of relative prices (embodied in the PPPs) has changed”. So there is no direct way in comparing iPLs of different PPP benchmark years, they simply belong to different price and quantities constellations.

As briefly mentioned above, an obvious, but nevertheless important, observation for the role of CPIs and each PPP update has been made by Klasen (2009). Comparing the changes between using different PPP benchmark years, and given that the available HHS are the same, the transition between PPPs affects the levels and hardly the trends. This is because the trends are mostly dictated by the application of CPIs in domestic terms.

At the same time, with respect to the CPI composition per se, Pogge (2013) convincingly argues that national CPI “is influenced most by the commodities on which most is spent.” Which means in turn that a CPI is a plutocratic index in this sense, as those who spend more influence the index more. Arguably those products that are not consumed by the people living in conditions of poverty should not be part of a proper price index for tracing poverty. And the weights used to construct such a price index "for the poor" should be representative of the consumption habits of those same individuals. Typically this is not the case for the CPIs applied in the global poverty literature.

1.6 Errors in PPP

Although the field of economics holds statistical significance dear, in global poverty research there seems to be little motivation in estimating, or disclosing, confidence intervals of the estimated poverty rates. As discussed above the PPP exchange rates are hardly a set of flawless figures. And this observation gains in importance when one uses such PPP rates for poverty estimates due to the methodological mismatches already mentioned.

The ICP PPP rounds do not report any confidence intervals of their PPP rates. In addition, by the construction of the PPP rates, any measurement errors in one country have an impact in the PPP estimate of all other countries. In that sense those errors are contagious, since they influence the entire series of poverty estimates for one specific country, and to a certain extend all the other countries as well. The size of those error terms is not marginal. For example, Deaton and Heston (2010) argues that the PPP estimates between China and USA contain a 25% error margin.

Along these lines, Deaton (2001, p.129) argues that since world primary commodity prices are “notoriously volatile”, while at the same time there are some countries for which those commodities consist a large part of their GDP, PPP exchange rates can vary considerably based on the world price of those

---

13 This interestingly translates to that, given the domestic real growth level in consumption, the MDG1 goal remains linked to the CPI application, and PPPs play a rather secondary and indirect role. The PPP role in trends could be attributed to shifting the level of iPL that may point it to a section of the domestic consumption distribution that may well differ in its steepness.

14 Without the underlying ICP data only rough approximations can be made by independent scholars Deaton and Dupriez (2011) Deaton (2012).
commodities in the base year of the ICP round. He further argues that this may explain the sensitivity of the African and Latin American with every other round of PPP estimates. This indicates the uncertainty and the volatility behind PPP estimates, or their use in years outside of the base year.

Deaton and Dupriez (2009) tabulate their "PPP for the poor estimates" (P4), and estimate the level of uncertainty between various reasonable choices of iPLs and P4 estimates. However, even if uncertainties from the use of various index numbers are considered by Deaton and Dupriez, no price measurement error, or an error capturing misrepresentation of the populations is included in their approach (ibid, p.40).

Finally, Ravallion (1994) investigates the impact of PPP measurement errors in the poverty ranking of countries. He shows that when those errors are identically distributed in each country then it can be shown that these errors do not affect the rankings. However, if the errors are heterogeneously distributed between countries then these rankings are not a priori robust to those PPP measurement errors.

2 Conclusions

The impact of PPP exchange rates in global poverty measurement cannot be overstated. Those rates make the standard approach to global poverty possible, by providing the empirical apparatus for creating the dollar-a-day approach. However, the important methodological mismatches between their specifications and its application to global poverty measurement, that are discussed above, provide good reasons for investigating alternatives as recommended by a number of scholars, such as Atkinson (2016). The cost of basic needs alternative method entirely avoids the use of PPP rates in estimating global poverty as discussed by Allen (2017). Moatsos (2017).

References


