Understanding PPPs and PPP-based National Accounts: Comment[†]

By Martin Ravallion*

While the exchange rates for purchasing power parities (PPP) are widely used, it is a fair guess that most users have never looked closely at how they are constructed and may thus have a rather limited understanding of them. That is in part because it has been hard to find a succinct and clear yet critical guide to PPPs, written for economists.

Angus Deaton and Alan Heston (2010) have done a valuable service in providing just such a guide. Their paper is also a timely contribution given that the final results of the 2005 International Comparison Program (ICP) were released in 2008, providing PPP exchange rates based on the ICP price surveys (World Bank 2008a, 2008b). And, a new version of Penn World Tables (PWT 7.0), using the price data from the 2005 ICP, is expected soon. The paper is also timely from the point of view of the design of the forthcoming 2011 ICP.

The authors generally focus on the methodological issues rather than specific country results. The big exception is China. The new PPP for China has attracted much public attention and Deaton and Heston (2010) refer to China 40 times. I will begin with some comments on what they say about China's PPP. I will then turn to the broader methodological issues they raise. The final section comments on some implications for current knowledge about macroeconomic aggregates and for cross-country regressions, including growth empirics.

I. China's Controversial PPP

The year 2005 was the first time China participated in the ICP, with past estimates based on old (1986) non-ICP price data for China and the United States, updated for differential inflation between the two countries. It would not be too much of an exaggeration to say that we have all been pretty much guessing China's PPP, and hence the real size of its economy, and it looks like the guesses were well off the mark. The 2005 ICP implies that China's gross domestic product (GDP) per capita at PPP for 2005 was $4,091, as compared to the prior estimate of $6,760 (World Bank 2008a)—a 40 percent drop in the estimate for the same year. This stems from a higher PPP. The 2005 ICP implies that China's "price level index"—PPP for GDP divided by market exchange rate (US=100)—is more than double the estimate used for 1993, which was around 20 percent (World Bank 2008a).

*Development Research Group, World Bank, 1818 H Street NW, Washington, DC, 20433 (mravallion@worldbank.org). The views expressed herein are those of the author and should not be attributed to the World Bank or any affiliated organization.

The size of the revision to China's GDP has led some observers to question whether the new PPP is believable. Deaton and Heston (2010) provide a seemingly striking calculation, based on an article by Surjit S. Bhalla (2008). Using the official growth rates, they project China's GDP per capita back to 1952 based on the new PPP, and calculate that the country's annual GDP then was $279 at 2005 international PPP dollars. This is deemed to be so low that, as Deaton and Heston (2010) state, "it is simply not possible that *both* the current PPP estimate of Chinese GDP and the official growth rates are correct." The authors appear to endorse Bhalla's view that the World Bank's new PPP for China implies that "most Asians (were) dead in 1950."

However, on closer scrutiny this troubling calculation is not quite so troubling. In fact we don't need to conclude that there is something wrong with either the PPPs or the growth rates to reconcile the observations made by Bhalla (2008) and Deaton and Heston (2010); all we need to believe is that China in 1952 was as poor as the poorest country *today*, which is the Democratic Republic of the Congo, with GDP per capita of $264 at 2005 PPP (World Bank 2008a). Many people survive on less than $279 per person, per year at 2005 PPP. Granted they live under wretched material conditions almost unknown in the United States or Western Europe, but they can still buy a frugal diet capable of supporting life (bulked up with starchy staples, though still with some vegetables, pulses, and fruit, and very modest amounts of meat, fish, milk, eggs, edible oil, and spices, with maybe $0.25 a day at PPP left for essential non-food items). And, it is not just a few rugged souls who live this way. From *PovcalNet* (combining the 2005 ICP results with household survey data), one can estimate that in 1981 over one billion people in the world lived below $279 per year.[1] And, many people are this poor even today. The number *living* below $279 was over 400 million in 2005 (100 million of them live in India, and about 40 million in China).

While one may not be convinced by the Bhalla-Deaton-Heston calculation, there are more compelling reasons to question China's data. On studying how China constructs its national accounts, Angus Maddison (1998) and Alwyn Young (2003) argue that the growth rate is overestimated, though there is scope for debate on how much so.[2] The PPP can also be questioned. On looking more closely at the design of the ICP price surveys for China, Shaohua Chen and Ravallion (2010a) argue that there is an upward bias in the price level due to undersampling of rural areas, where prices are lower; their correction for this bias still implies a substantially higher price level in China than was thought. Yes, there are problems with China's data. But, we can only know that from careful scrutiny of those data.

There is also an economic reason why we should not be surprised that China's PPP has been rising relative to its market exchange rate. For essentially the same reason that one uses PPPs rather than market exchange rates—namely the Balassa-Samuelson effect—rising productivity in China's tradable goods sector can be expected to put upward pressure on the overall price level. Elsewhere, I have

---

[1] Even allowing mean consumption to be (say) 25 percent below GDP per capita, the count is almost 600 million.

[2] For example, see Carston A. Holz's (2006) comments on Maddison's (1998) assumptions.

estimated that this "dynamic Penn effect" accounts for about two-thirds of the rise in China's price level index between 1993 and 2005 (Ravallion 2010).

II. Debates on PPP Methodology

There are many ways one can go from the ICP's raw price data by commodities and countries to PPP exchange rates. On top of the standard issues of bilateral price indices, a PPP is a multilateral index, with one value (for GDP or each of its components) for each country. In aggregating the prices to form a multilateral index, PWT has used the Geary-Khamis (GK) method while the World Bank has used the Elteto-Kölves-Szulc (EKS) method; Deaton and Heston (2010) explain the differences well.

The choice depends to some degree on what the PPP is used for. Deaton and Heston argue that the GK method is better for national-income accounting, given that it guarantees exact aggregation across GDP components, which the EKS method does not. They do not give examples of applications that require exact aggregation, though one can think of some, including cross-country comparative studies of the shares of GDP, such as in studying aggregate savings or structural transformation in development. However, exact aggregation does not appear to be required for many macroeconomic applications of PPPs.[3]

An important class of applications for PPPs is in international welfare comparisons (including in global poverty measurement). For that purpose, many economists (myself included) will be more comfortable with the EKS method, given that in binary comparisons it collapses to the familiar Fisher Index, which provides a second-order approximation to a utility-consistent ("true") cost-of-living index for homothetic preferences. Of course, one can question these assumptions. But at least this method can claim welfare consistency in special cases, which the GK method cannot. It is also well known that the GK index overstates real incomes in poor countries (given that the international prices are quantity weighted, though this can be fixed).

This choice goes to the heart of what we mean by a "PPP." At the outset of their paper, Deaton and Heston (2010) define it as a measure of "how much local currency is needed to buy as much as does the currency in the numeraire country." Given that PPPs are widely used for international welfare comparisons, a better definition might be "how much local currency is needed to attain the same level of welfare as can be reached with one unit of the numeraire currency." The issue is then defining a consistent concept of welfare. If one assumes a common utility function (with standard properties) defined on the consumption vector then this usefully constrains the set of options for deriving a PPP. Deaton and Heston (2010) are critical of those assumptions on empirical grounds and (hence) question the usefulness of welfare-based, cost-of-living indices. However, I am not convinced that assuming common preferences is as unreasonable as the authors contend. Revealed preference

---

[3] Of course, exact aggregation of GDP components is often required for macroeconomic analysis at country level, but such analysis does not typically require PPPs.

tests using ICP data have actually been quite supportive of the common-preferences assumption, though homotheticity is more problematic.[4]

Furthermore, even if we grant that these assumptions do not always hold, that does not make it unreasonable to impose them in making welfare comparisons. Indeed, it is hard to see how any progress can be made in the task of making consistent welfare comparisons if one does not impose some sort of common preference structure. And, whether one likes it or not, PPPs will be used in making international *welfare* comparisons. The fact that preferences differ somewhat in reality or are not homothetic does not rule out the case for the economic approach to welfare comparisons. The debate on this issue will no doubt continue.

Deaton and Heston (2010) pay (welcome) attention to various problems related to the underlying price data that enter into the aggregation formulae, to derive a PPP. An important issue is setting quality standards. Prior to the 2005 ICP, the standards used in defining the commodity to be priced in each country appear to have been rather weak. Clearly this can be a serious concern for welfare comparisons (such as in measuring poverty or inequality), given that there is likely to be an income effect on the quality of commodities consumed. Without strict standards in defining the products to be priced, there is a risk that one will underestimate the cost of living in poor countries by confusing quality differences with price differences. PPPs will be underestimated in poor countries. The only way to reduce this "quality-bias" is through detailed product listings and descriptions, which add significantly to the cost of data collection. The better-funded 2005 round of the ICP, managed by the World Bank, made considerable effort to address the quality-bias problem and this was a major improvement in the 2005 ICP, as Deaton and Heston (2010) acknowledge.

There are also problems in aggregating individual prices *within* the ICP's 115 basic headings. By definition, there are no expenditure weights available from the national accounts below the basic headings, making this a problem of constructing what are called elementary price indices. It is clearly desirable that at least some of the goods in the bundles specified for a given region are commonly consumed ("representative") in all countries. But naturally this will not be true of all goods. There will be missing values, which are usually filled in by the "country-product dummy" (CPD) method described by Deaton and Heston (2010). There are also "nonrepresentative" goods that can only be found in specialty outlets at (probably) high prices. Deaton and Heston (2010) point out that rice is hard to find in Ethiopia and teff (the staple in Ethiopia) is hard to find in Thailand. This may be rather forced as an example, given that rice is available in Ethiopia and *teff* is not in the Asian region's bundle of goods for the ICP, but the point is well taken in principle. The concern is that the "specialty-store prices" may well have undue weight within a given basic heading.

---

[4] Robert Ackland, Steve Dowrick, and Benoit Freyens (2007) apply revealed preference tests to the 6,555 bilateral comparisons that are possible amongst the 115 ICP 1993 countries and find only one pair for which common preferences assumption could be rejected, namely Armenia and Uzbekistan. However, the assumption of common *homothetic* preferences fails their tests for 30 percent of countries.

This is a long-standing issue.[5] One needs a common reference bundle but some of the goods in that bundle may not be representative of consumption patterns in some countries. One possible response is to only collect the prices of seemingly "representative" goods. But, this would reintroduce the quality bias that has plagued past ICP rounds, and can lead to substantial underestimation of PPPs in poor countries. Using region-specific bundles helps get around this problem. In the 2005 ICP, each regional office started from scratch to develop its own bundle, working with the countries concerned. However, the problem reappears in the ring comparisons used to link PPPs across regions through a globally common set of goods priced in a subset of "ring countries" spanning all regions. (There were 18 ring countries in the 2005 ICP.) We do not know how serious a problem this is in practice.

Another issue concerns the "regional fixity" assumption built into the World Bank's ring comparisons. Essentially, PPPs are first estimated *within* each region and then linked globally through the global prices collected for the set of ring countries. In doing so, the original PPPs' relativities within each region are preserved. Deaton and Heston (2010) point out that this is not the only option, and nor is it obviously the best option. The aggregation procedure privileges the original price relativities, essentially ignoring the new ring prices beyond their role in estimating the region-specific factors. The authors conjecture at one point that the fixity assumption is largely political, rather than statistical, in motivation. Yet their own concerns about the representativeness of prices—the "specialty-store" problem— would surely suggest that one would not want to use the prices collected in the ring countries for anything more than the task of estimating the region effect. So there does appear to be a statistical defense for this methodological choice in estimating PPPs.

III. Some Implications

The data and methodological issues underlying PPPs have two important implications for applied work. First, they speak to the need for caution in making PPP comparisons across ICP rounds, or between the World Bank's PPPs and those in PWT. A common practice in data construction and research is to only do the PPP conversion at one date—the most recent ICP round—and then rely on national data sources for inter-temporal comparisons. There are arguments for and against this practice. As I have already noted, the Balassa-Samuelson effect should be operating within growing developing countries (as well as between developed and developing countries). Then intertemporal comparisons may cease to be consistent with international prices, as Simon Johnson et al. (2009) have recently warned. Against this view, it has been argued that consistency with national prices trumps international prices for some purposes, including growth empirics (as argued by Daniel A. Nuxoll 1996, amongst others). However, whichever position one takes on this issue, users of PPPs should be aware of the data comparability and measurement error problems.

---

[5] J. R. Cuthbert and M. Cuthbert (1988) provide a full discussion of the issues in aggregating prices to the basic heading level, and study the sensitivity of PPPs to various methods.

Second, the comparability problems speak to the need to revisit past research findings in the light of each new round of the ICP, particularly when those findings rely heavily on developing-country data. We already know some of the implications for current knowledge. Inequality in the world as a whole is even higher than we thought (Branko Milanovic 2009). The level of poverty at any given international poverty line in PPP dollars is also higher, though the PPP revisions imply that the national poverty lines in poor countries are lower than we thought (Chen and Ravallion 2010b).

The literature using cross-country regressions contains a vast array of other PPP applications. GDP at PPP is widely used as a control variable in cross-country regressions, where the dependent variables range widely from GDP growth rates to measures of inequality and of social and political development. The implications for empirical research using such cross-country regressions are less clear on a priori grounds. There is the usual attenuation bias due to the measurement errors; for example, in standard tests for conditional convergence we expect the true speed of convergence to be even higher than implied by the OLS regression coefficient. However, there is evidence of a more specific structure to the PPP revisions implied by the 2005 ICP, namely that the upward revision to the 1993 PPP implied by the 2005 ICP tends to be larger for poorer countries (at least amongst developing countries).[6] This implies an offsetting effect to the attenuation bias; indeed, if this "structure effect" is strong enough, then the true speed of convergence may be lower than we thought. That remains to be investigated.

There are also important implications for public access to ICP data, and the ICP's reporting of results. On access, it is evident that many of the sensitivity tests and alternative PPP calculations that one would like to do require the micro price data (the underlying prices, not just the basic heading average prices). For example, to control for the type of outlet in CPD regressions (to test for the claimed "specialty store" effect emphasized by Deaton and Heston 2010), or to test the effect of relaxing fixity, one would need the micro data. Only then will we know how important are the concerns raised by Deaton, Heston, and others. Most readers (including myself) will agree with Deaton and Heston's (2010) plea to make the underlying price data available to researchers. Yet the ICP has not allowed such access (even to their colleagues within the World Bank) on the grounds that they need to defend the "confidentiality of price data." Yet confidentiality should not be such a problem given that prices for most goods and services are essentially public knowledge at country level.

Future ICP (and national accounts) reporting should provide more prominent "quality tags" to each (country-specific) point estimate, such as whether or not the country in question actually participated in the ICP price surveys, whether there are any likely biases in the ICP sampling of outlets, and information on the distribution by type of outlet. Better data on the degree of imprecision in widely-held priors based on macroeconomic data would ultimately enhance public confidence in those

---

[6] This is demonstrated by Ravallion, Chen, and Prem Sangraula (2009).

data, help explain otherwise puzzling data revisions, and could also facilitate more robust analyses of those data.

IV. Conclusion

Large data revisions are never easy to absorb. But they are often needed, notably when priors had been formed on weak data. We have learned a great deal from the various rounds of the ICP, and the quality of the data has improved. Much more could be learned, including at sub-national level, provided that the underlying (micro) price data become publicly available. Research on the data, and on the methodological issues underlying PPP estimation, will hopefully continue, as will improvements in the data and methods. Deaton and Heston's (2010) masterful overview of the issues provides a foundation for that future work.

REFERENCES

Ackland, Robert, Steve Dowrick, and Benoit Freyens. 2007. "Measuring Global Poverty: Why PPP Methods Matter." Unpublished.

Bhalla, Surjit S. 2008. "World Bank—Most Asians Dead in 1950." *Business Standard*, August 23.

Chen, Shaohua, and Martin Ravallion. 2010a. "China Is Poorer than We Thought, but No Less Successful in the Fight against Poverty." In *Debates on the Measurement of Global Poverty*, ed. Sudhir Anand, Paul Segal, and Joseph E. Stiglitz, 327–40. New York: Oxford University Press.

Chen, Shaohua, and Martin Ravallion. 2010b. "The Developing World is Poorer Than We Thought, but No Less Successful in the Fight against Poverty." *Quarterly Journal of Economics* (on press).

Cuthbert, J. R., and M. Cuthbert. 1988. "On Aggregation Methods of Purchasing Power Parities." Organisation for Economic Co-operation and Development (OECD) Department of Economics and Statistics Working Paper 56.

Deaton, Angus, and Alan Heston. 2010. "Understanding PPPs and PPP Based National Accounts." *American Economic Journal: Macroeconomics*, 2(4): 1–35.

Holz, Carsten A. 2006. "China's Reform Period Economic Growth: How Reliable Are Angus Maddison's Estimates?" *Review of Income and Wealth*, 52(1): 85–119.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2009. "Is Newer Better? The Penn World Table Revisions and Their Impact on Growth Estimates." National Bureau of Economic Research Working Paper 15455.

Maddison, Angus. 1998. *Chinese Economic Performance in the Long Run*. Paris: Organisation for Economic Co-operation and Development (OECD).

Milanovic, Branko. 2009. "Global Inequality Recalculated: The Effect of New 2005 PPP Estimates on Global Inequality." World Bank Policy Research Working Paper 5061.

Nuxoll, Daniel A. 1994. "Differences in Relative Prices and International Differences in Growth Rates." *American Economic Review*, 84(5): 1423–36.

Ravallion, Martin. 2010. "Price Levels and Economic Growth: Making Sense of the PPP Changes between ICP Rounds." World Bank Policy Research Working Paper 5229.

Ravallion, Martin, Shaohua Chen, and Prem Sangraula. 2009. "Dollar a Day Revisited." *World Bank Economic Review*, 23(2): 163–84.

World Bank. 2008a. *Comparison of New 2005 PPPs with Previous Estimates: Appendix G Revised: Global Purchasing Power Parities and Real Expenditures.* Washington, DC.

World Bank. 2008b. *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program.* International Comparison Program and International Bank for Reconstruction and Development. Washington, DC.

Young, Alwyn. 2003. "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period." *Journal of Political Economy*, 111(6): 1220–61.