Shortcomings of and Improvements to Measures of Income across Countries

By Julieta Caunedo and Riccardo DiCecio

Every man is rich or poor according to the degree in which he can afford to enjoy the necessaries, conveniencies, and amusements of human life.¹

—Adam Smith

The task of building measures of Gross Domestic Product (GDP) that allow for comparing standards of living across countries presents several challenges. In addition, data revisions can have surprising effects. Consider two examples:

• The 2010 version of the World Bank’s World Development Indicators (WDI) implies that the United States was 10 times richer than China in 2005; the previous version (2007) implied that the United States was six times richer than China for the same year. Also for 2005, India was 12 times poorer than the United States in the first version of the WDI and 18 times poorer in the latest version.

• A popular source of real GDP data used in countless studies, the Penn World Table (PWT),² is not free of inconsistencies either. For example, differences between the latest two versions—both covering data for the year 1996—reach a standard deviation of 7.7 percent in annual growth rates for countries in the bottom third of the income distribution.³

These discrepancies are relevant for policy decisions. For example, the European Commission uses GDP per capita, adjusted for purchasing power parity (PPP), in deciding how to allot structural funds; these funds—25 percent of the EC’s total budget—are used to smooth disparities between and within member states.⁴

Also, assessing the success of policies designed to fight extreme poverty across the world depends on the measure used to define the poverty line.⁵ For example, when the World Bank decided in August 2008 that the official poverty threshold would rise from $1.08 of income a day to $1.25, an additional 430 million people around the world were automatically classified as being impoverished.

Comparable Measures of Output: Diagnosis

There are alternative ways to measure output in an economy: adding up the value added in each sector of the economy (production approach) or adding the value of total expenditure, i.e., consumption, investment, government spending and net purchases from abroad (or current account). Most of the national accounting is done using the latter.

One obvious difficulty in comparing income across countries stems from the fact that different countries use different currencies. The use of official exchange rates would not provide an adequate comparison. For example, if the Mexican peso were to depreciate by 10 percent with respect to the dollar, the GDP of Mexico would fall by the same amount when measured in dollars. However, if prices and incomes in Mexico were unchanged, Mexican residents would not be poorer by 10 percent.⁶ The Big Mac Index constructed by The Economist gives us a better comparison. As of July 2010, we can buy a Big Mac for $3.73 on average in the U.S. and for 32 pesos in Mexico. The burger exchange rate is 32/3.73=8.57 pesos per dollar. At such an exchange rate, a burger in Mexico and in the U.S. would have the same price in dollars.⁷ However, the actual nominal exchange rate is roughly 13 pesos per dollar: The dollars necessary to buy a burger in Mexico are not enough to buy the same burger in the U.S. This is what in economics jargon is called the purchasing power parity (PPP) adjustment. Still, moving from what theory suggests as the correct measure to the actual estimations is not without controversy. We wish people were to consume Big Macs only!

Some of the main issues in constructing these measures are:

1. People in different countries typically consume different baskets of goods. For example, the per capita consumption of meat in Argentina is about 70 times larger than in India, where cow meat is not usually part of the diet. However, price indices that allow for international comparisons should be pricing the same basket of goods.

2. Even if the bundle is the same, its value should be computed using relative prices across countries (multilateral indexes). In general, durable goods in terms of consumption goods are more expensive in developing countries than they are in the developed world, and, vice versa, services are relatively cheaper in developing countries. The PWT uses a valuation of goods that tends to overstate the value of consumption in poor countries.

3. It is difficult to value activities related to the service sector (e.g., housing rental, government services, health care): What is the value added to the economy of a teacher?

4. Measures of real GDP that are based on...
As the true GDP growth ranges from –0.25 official figures and what the authors claim by WDI. The overall difference between the 10-year growth rate for Tajikistan was –0.06 authors of this study found that the “true” GDP growth.9 For example, the satellite images, to obtain a better estimate the actual growth of some countries. A recent paper develops a framework that combines measured GDP growth with the components. 

More orthodox attempts aim at solving the problem of comparable bundles of goods. The latest PPP measures are built upon regional data, which typically compare groups of countries with similar economic structures and consumption patterns. Then, a few countries are selected as “bridges” to allow for cross-regional comparisons. An issue with this methodology is that the relative ranking of economies by GDP per capita may depend on the composition of the group of economies being compared.10

As for the treatment of the net foreign balance, some authors point out the importance of distinguishing the expenditures approach from the production approach to construct real GDP.11 Real GDP constructed from the production side measures the production possibilities of an economy and should not take the terms of trade into account. Even though real GDP data in the PWT are constructed according to the expenditure approach, the growth rates are more similar to those of production-based real GDP. For a sample of 151 countries, the aforementioned authors found that for one-third of them, expenditure-based real GDP is above output-based real GDP. When assessing how rich are the rich, complementing current measures with output-based series may improve the quality of the analysis.

expenditure—the International Comparison Program (ICP) and PWT—are highly influenced by the relative price of the country’s imports and exports, the so-called terms of trade. These measures tend to overstate physical output in countries that face a high relative price of exports.8

5. When aggregating data, it is common practice to use fixed shares of consumption, investment and public expenditure (the one corresponding to some arbitrary base year). This is problematic because changing base years (and, therefore, the contribution of each item in total output) may induce movements in estimates that do not stem from any fundamental change in value of the components.

Improving Matters

In view of these limitations, economists have relied on ingenious measures to approximate the actual growth of some countries. A recent paper develops a framework that combines measured GDP growth with growth in lights on earth, as measured from satellite images, to obtain a better estimate of “true” GDP growth.9 For example, the authors of this study found that the “true” 10-year growth rate for Tajikistan was –0.06 percent instead of –0.227 percent as reported by WDI. The overall difference between the official figures and what the authors claim as the true GDP growth ranges from –0.25 percent to 0.25 percent.
How Much Do We Actually Know?

It is quite unrealistic to believe that the comparisons between poor and rich countries are so far off that the relative position of countries would be reversed. However, the picture gets blurry when looking at the poorest economies. The figure depicts the change in ranking of Eastern African countries due to the WDI update. Countries are ranked from poorer to richer (left to right) in 2005 based on the latest version of the WDI (2010). The ranking gets shuffled if one uses the 2005 figures from the previous version of the WDI (2007).

Although a 40 percent margin of error is allowed for countries with the lowest data quality in the PWT, it is not plausible to attribute all of the inconsistencies to poor data quality. Merely changing the base year creates standard deviations in the differences of annual growth rates as large as 5.4 percent on average.

Moreover, the current measures tend to build price and quantity indices for baskets of goods resembling more those consumed in the rich than in the poorest economies. Arguably, the most promising project directed to partially solve this problem seems to rely on the regional grouping of countries.12

We also expect the issue with the treatment of international accounts to be solved soon. If not, the user should be particularly careful when looking at countries that are resource-rich or that have an ample exportable base in commodities. These countries are the ones most affected by changes in relative prices of their exportable goods.

We should expect further adjustments in the growth figures across countries. Hopefully, adjustment in levels and growth rates will be smoothed along time. Common sense remains the best way to assess results. Robustness checking should be combined with in-depth understanding of how data are constructed.13

Riccardo DiCecio is an economist at the Federal Reserve Bank of St. Louis. Julieta Caunedo is a research analyst. Go to http://research.stlouisfed.org/econ/dicecio/ to see more of DiCecio’s work.