Purchasing power parity measures: advantages and limitations

By Samuel Moon, Sheila Page,
Barry Rodin and Donald Roy

Discussion and analysis on international comparison using Purchasing Power Parity (PPP) has been somewhat neglected in recent years. With recent shifts in economic growth and constantly changing poverty dynamics, as well as forthcoming major evaluations of progress on the Millennium Development Goals (MDGs), it is an opportune time to revive the debate. This paper examines what PPP tries to do, some of the problems of measuring it and using different measures, and why it matters.

Why we need PPP

Policy-makers, researchers, businesses and consumers all want to compare incomes and spending, often when prices are different or changing. The comparison of incomes and measurement of changes in incomes are essential tools in analysing the success of economic policies and, in particular, in asking questions about development. Observers have a wide range of purposes for such comparisons, so there is no single answer to what is the 'right' measure of prices to use, and problems in the data and methodology mean that there may be no single answer even for a specific question. This paper explores how estimates of incomes and prices can be calculated and gives examples of why choosing the best indicator matters. It surveys a range of different purposes for comparing incomes and suggests some of the decisions that must be made in choosing how to make different types of comparison. In short, users must define clearly what they want to know, what they already know or can assume about

how the people whose incomes are being measured behave, and the trade-off between the appropriateness of the measure and the costs of improving it.

Starting in the 1990s, the major restructuring of production in eastern Europe, the adoption of numerical targets in development (reducing 'less than a dollar a day' poverty), and the growing economic and political power of countries normally considered 'poor' (especially China) have increased interest in this subject. New and revised international data from the 2005 round of the International Comparison Programme (ICP) published in 2008 meant that estimates of economic activity measured at purchasing power parity (PPP) are now available for the vast majority of countries, including China for the first time. This has led to new comparisons, but also to growing awareness of the limitations of the data.

> **Box 1**
>
> A **consumer price index**, CPI, is a tool to measure changes in prices of a set basket of goods over time, or differences in the prices of that same basket of goods in different regions.
>
> Consumer preferences are captured using a **household survey** that measures expenditure patterns over different demographics.
>
> **Purchasing Power Parity**, PPP, is a mechanism for accounting for different relative costs of goods when undertaking comparative analysis of expenditures and incomes in different countries.

Measuring prices

When comparing incomes across time or across countries or even within countries, the first set of choices to be made is about which prices go into the raw data. Goods and services vary in quality; prices often can vary drastically by season, but in different ways in different parts of a country; goods can be bought in different quantities from different suppliers in different locations (and all these may be correlated with the buyers' incomes); they are bought in different proportions by different buyers (including by the different groups to be compared); the access to markets often varies for different demographics of the population for a variety of reasons. There are, therefore, both price differences for the same commodity or service and differences in the importance of the same commodity for different consumers, within a country and between countries.

The data require detailed price and household surveys to determine what the group of interest buys, and where and how they buy it. How detailed do they need to be? This is partly a matter of judgement, to be revisited as more information becomes available (how large are the differences, so how many different types of observation are needed?) and partly of cost.

Developing consumer price indices for a comprehensive analysis or even for a particular group in a given country may be extremely difficult and/or costly. To give an example, there is often a focus on determining the income dynamics of the rural poor. But Consumer Price Surveys are often only taken at urban trading centres. While the rural population may have to go into these trading centres to access some goods, they will be exposed to additional (and highly variable) costs of transport and market knowledge. For other goods, the rural population may largely produce their own food supplies, and 'farm gate' food prices are not likely to be measured adequately by a Consumer Price Index (CPI) observed at trading centres. Preferences differ too. One region, tribe or gender, for example, may consistently choose to consume a basket of food goods that diverges from the national average.

In Uganda, the national poverty line is calculated with a basket of goods that gives a heavy weighting to matooke, a type of green banana that is traditional in much of the central area of the country, but is relatively more expensive and less calorie efficient than other staples. Most of the population elsewhere in the country prefer staples that are naturally cheaper per calorie, and therefore require a lower overall income to afford the basket of food providing the same calorie levels. This may have systemic effects on over-estimating poverty in areas that do not favour matooke, but the example raises questions about the definition of poverty, which is addressed later in this paper.

In order to compare incomes between two groups with different spending patterns, some method of adjusting these must be found. This can be done by calculating an average (for which there are different methods) or by doing two comparisons, using each group's pattern. These will give different answers, and can give different rankings. The problems become more complex when there are more than two groups (for example all the countries in the world) and even more complex if various levels of analysis are needed (for example, comparing members of a group such as the European Union among each other and then, as a group, with the rest of the world).

For comparisons within one group, with relatively small changes in price, relatively small changes in income can be measured. For different populations with different compositions of spending any comparison will vary depending on the method of accounting for the differences. For large differences or changes in price, assumptions must be made about how substitution from a high priced good to another affects the comparison: substitution will lower the prices paid, apparently increasing income, but will reduce the perceived real income because the consumer preferred the higher price good.

Examples of using income comparisons to answer development questions

Which countries have the highest or lowest total national incomes?

If this is a question about economic power, then spending power within each country may not be relevant, so price comparisons are unlikely to be needed.

But if it is a question about which countries have achieved most or which should contribute most to global needs, whether formal contributions to international organisations or through aid, security spending, etc., people may want an answer based on the level of per capita incomes, adjusted for different prices, i.e. PPP calculations.

One topical reason for such comparisons is to identify how large current differences in income levels are in order to identify how the contribution of currently low-income countries to climate change will increase as their incomes 'converge' on those of high-income countries. The Intergovernmental Panel on Climate Change (IPCC) has used estimates of economic activity based on market exchange rates rather than on purchasing power parities. These overstate both significantly and systematically the gap in wealth between rich and poor countries and thus the amount of economic growth required to achieve convergence over the current century. PPP estimates of economic activity would form a much sounder basis for economic projections, and in this

Case Study 1: Use of cost of living indices/PPPs when managing international assignments

In the past 20 years the international mobility of staff has been seen as vital for the success of many organisations' global operations. A key consideration when resourcing these activities is assessing pay for expatriates as cost effectively as possible.

Although there are many criteria and objectives to consider when formulating pay methods for international assignees, the most common approach ensures that pay is at least equitable or consistent with their home country peers. In other words, home country living standards should be protected for economic variables such as differences in living costs and levels of taxation, which still exist even in Western Europe.

One important tool to help protect home country living standards is the Purchasing Power Parity (PPP) or cost of living index, which gives the difference in living costs between two locations, based on a specific set of assumptions. The proportion of the assignee's basic or notional home salary spent on day to day goods and services (spendable income) is adjusted by the PPP to determine the necessary cost of living allowance to protect specified living standards. Tables, normally produced by data suppliers or consultants, give spendable income proportions for specific incomes and family sizes. These are normally derived from household expenditure surveys published by official statistical offices. There are several key issues to consider when calculating PPPs.

Goods and services selected: these typically cover items of day to day consumption by the expatriate family whilst on assignment. These often exclude items for which the cost is covered separately by the employer (e.g. host country housing), or provided as benefits in kind (e.g. children's education, medical insurance, company cars etc.). Item specifications need to be as unambiguous as possible to ensure prices are selected for the same acceptable quality in all locations surveyed. Items chosen for the shopping basket are those that are typically purchased in the assignee's home country but that are also available as far as possible in all locations. This also helps ensure consistent comparisons between locations.

Weighting of items used in the PPP/index calculations: these normally reflect the income group or professional status of assignees (e.g. middle to upper incomes or skilled/professional/managerial levels) and specified family size (e.g. married couple with one or two children). Latest available government statistics on household family expenditure, for specified income groups and family sizes, are generally referred to when determining item weights.

Pricing of items: these should be from shops and outlets, providing goods and services of comparable quality to those in the home country. About 50% of employers using PPP/index calculations assume that shopping abroad is not as cost effective as at home ('green' shopper). The remainder use figures based on the premise that expatriates shop with the same level of cost effectiveness in both their home and assignment country.

Types of indices: indices based on the assignee's home country consumption pattern are non-reversible. The cost of living differential for an assignment from A to B is not the same as for assignments in the reverse direction from B to A; a different weighting pattern, based on the home country of each assignee, is used in the calculations. Another approach is to calculate indices based on a combined international or regional spending pattern. It is assumed that assignees shop at an equal level of cost effectiveness in all locations, including at home. This type of index is both reversible and transitive (i.e. differences between locations remain the same, regardless of the assignee's base country). This index is applied by a minority of organisations who assign many different nationalities to and from many locations and expect their assignees to adapt to an international or 'nomadic' lifestyle.

Although cost of living allowances may represent a relatively small proportion of the total cost of an international assignment, they often attract the most discussion and questions from expatriates. Quoting one Human Resources practitioner responsible for expatriate pay calculations and negotiations: 'everybody is an expert on the cost of living and weather'. Consequently it is essential to ensure that indices are based on sound and defensible criteria, and are calculated using rigorous statistical and validation processes to ensure both accuracy and consistency.

Case Study 2: Wanless in context: International comparisons of real expenditure on health care in developed countries

This study looks at how spending on health care, using national accounts definitions, has been compared between the United Kingdom and other developed countries, based on the Wanless Report (2002). It suggests that the use by Wanless of general purchasing power parities rather than ones specific to health care led to significant distortions.

The Wanless Report referred specifically to real expenditure on health care in 1998 in eight countries as comparators with the United Kingdom (Table 1). Two of these were omitted from later analysis (Japan and the United States).

Table 1: Real expenditures per capita on health care in 1998 and 1999 using general and sector-specific PPP methods

Country	General PPP		Sector specific from 1999	
	in 1998 as quoted by Wanless	1999	EKS PPP	G-K PPP
Australia	2,085	1,790	2,177	1,904
Canada	2,360	1,836	2,234	1,766
France	2,034	1,900	2,425	1,958
Germany	2,367	1,751	2,159	1,903
Japan	1,795	1,732	2,399	2,138
Netherlands	2,150	1,468	2,265	1,984
New Zealand	1,440	1,380	1,583	1,283
Sweden	1,732	1,699	1,797	1,531
United Kingdom	1,510	1,519	1,938	1,489
United States	4,165	3,906	2,906	3,005

Source: 1998 figures taken from Wanless, Table 5.1, p.64; 1999 figures from OECD, 'Purchasing Power and Real Expenditures 1999 Benchmark Year'.

Wanless relied on estimates of real expenditure per head supplied by the OECD. These are available annually and are based on general purchasing power parity conversion. Health expenditure data converted using specific parities are available triennially as a by-product of overall national accounts comparisons undertaken by a different department in OECD. 1998, the year used by Wanless was not one covered by the triennial comparisons.

However 1999 was covered. Table 1 also shows 1999, using three alternative measures of real expenditure on health care. The first converts using the whole economy purchasing power parity and is comparable to the estimates in the first column for the previous year. The second and third use only health-specific prices. The second (Elteto-Koves-Szulc – EKS PPP) uses a specific parity derived from the method currently in favour with international organisations; the third (Geary-Khamis – G-K PPP) is derived from an earlier method.

The effect of using specific PPP in 1999 is to reduce the measured real expenditure in the United States and to alter the pattern elsewhere – whether real expenditure per head on health care in Sweden was greater or less than in the United Kingdom in 1999 appears to depend on the specific parity used (either way the Swedish figure was close enough to that for the United Kingdom to cast a shadow of doubt over the automatic assumption that increases in real expenditure were required to produce better results). This illustrates both the risks of using a general measure and the differences possible even between sector estimates.

Estimates using both specific parities are available for 2002 and estimates based on Elteto-Koves-Szulc for 2005 and are shown in Table 2, expressed in OECD dollars for 2002 and international dollars for 2005.

Table 2: Real expenditure per head on health care using sector-specific parities in 2002 and 2005

Country	EKS PPP 2002	G-KPPP 2002	EKS PPP 2005
Australia	3,019	2,442	3,436
Canada	2,539	1,765	3,270
France	3,350	2,310	3,934
Germany	2,588	2,097	4,124
Japan	2,863	2,334	4,653
Netherlands	2,615	2,223	3,681
New Zealand	2,201	1,598	2,720
Sweden	2,655	1,669	3,633
United Kingdom	2,607	1,672	3,661
United States	3,121	3,426	5,843

Source: EKS figures from OECD (2002: 134-137).

The figures for 2005 suggest that real expenditure on health care per head in the United Kingdom had overtaken that in Australia and Canada and was level with that in the Netherlands and Sweden and closing in on that of France.

Data on some of the comparator countries are available also for the first half of the 1970s (covered by the first three phases of the International Comparison Project/Programme (ICP)). These are in international dollars of 1970, 1973 and 1975, calculated using Geary-Khamis.

Table 3 shows real expenditure per head in the six countries expressed in international dollars of 1970, 1973 and 1975 respectively. Wanless-type estimates can be reconstructed quickly for the first two years, as also shown in the table.

Table 3: Real expenditure per head on health care in the 1970s

Country	Using sector-specific parities			Using whole economy parities	
	1970	1973	1975	1970	1973
France	212.0	309.1	386.6	216.2	315.3
Germany	186.7	236.3	400.4	196.0	295.4
Japan	209.1	311.4	335.3	123.4	174.4
Netherlands	170.5	223.5	251.6	156.2	236.9
United Kingdom	138.1	186.3	328.3	125.7	160.2
United States	175.6	253.9	401.3	316.1	434.2

Source: Authors' calculations from OECD (2002: 176-187).

Although the United Kingdom had the lowest level of real expenditure on health care in the group in 1970 and 1973, and thus in the figures published in 1978, the difference between the UK and France and Germany would have seemed much greater using the general parity comparison than on the actual ICP results. The Chief Economist at the then UK Department of Health and Social Security was aware of the difference between general and specific parities and was able to make appropriate and effective use of them (Roy, 1982). It is possible, therefore, and may be probable, that policy-makers in the late 1970s had superior analytical tools at their disposal than those used by their successors at the turn of the century.

case, would imply a much smaller increase in effects on the climate from catching up by developing countries.

Which countries have the highest or lowest per capita incomes?
If this is a question about standards of living, then this is the most obvious application for PPP measures, although with all the problems discussed above. The evidence of recent, increasingly detailed, disaggregation of spending and prices by income level is that prices within countries tend to be higher for the poor (who are less able to buy in large quantities or to find cheaper, but perhaps more distant, suppliers), so that estimates of incomes in low-income, relatively unequal countries are being reduced as measurement improves. (Estimates for both China and India have been lowered, for example.) This could slightly offset the differences in the gap between high- and low-income countries measured at market prices or PPP prices, and thus some of the reduction in estimates of the costs of convergence.

Measuring income differences reliably is also important for those who move between countries. As the first Case Study, by Barry Rodin, demonstrates, even if all prices and consumption patterns are known, there is no single answer to the question: what is the difference in prices for which an expatriate needs to be compensated?

Which countries' incomes have grown most or least?
Price changes within each country must be calculated accurately and with congruent allowances for differences in income, location, etc., but the prices and composition of spending do not need to be compared across countries. Nevertheless, the growing awareness of the need for disaggregating spending patterns and prices means that comparisons with older surveys become less reliable.

What level of per capita income measures 'poverty'?
Calculations of poverty lines are based on the costs of consuming goods considered essential to life, sometimes called 'the real cost of meeting basic human needs'. While they differ from measures based on actual spending patterns, they present the same problems of choosing which prices to use for each good: identifying the appropriate locations, types of market, and units of purchase. Comparisons across countries must take account of different 'basic' goods, depending on the staple food, type of clothing, needs for shelter, etc. This also brings up the problematic question of defining poverty. The definition of poverty will change by demographic

and over time. At its core, poverty measurement will revolve around having enough income to consume a minimum number of calories per day. There is also, however, a core set of non-food items that are needed, and the selection of these will affect the share of food and non-food costs in the minimum basket of goods. As countries develop, this non-food share will grow and begin to include costs of education, mobile phone usage, utilities, transport, fuel and so on. Thus both the choice of goods and the types of comparison are subject to debate.

Has a country (or region or the world) reduced poverty?
This has become a prominent question, as halving poverty is one of the MDGs. Once a poverty line and the goods to be included are chosen, national poverty can be measured and country comparison can be made by comparing each country's average income per capita to the poverty line or by calculating the proportion of each country's population falling below that line. The first method does not measure poverty well in countries with very unequal distribution. The second does not measure changes taking place below the line (moving people up, but not over the line), but does include even very small movements, if they are at the line. Any measure is subject to the problem that improving the measurement of the prices paid by poor people may generate an increase in measured poverty or income disparities.

What level of per capita income should be used to measure 'development'?
Per capita incomes are used to define 'Least Developed Countries' (LDCs), a legal, UN category used in defining UN and World Trade Organization obligations and by some donors, including the EU, to allocate trade concessions. Low-Income Countries (LICs) is a variable category, but most multilateral and national donor agencies use some definition of low income to define eligibility for aid. Other income categories are used to decide which countries should have full international obligations and which should have differential treatment in determining eligibility for aid or other special funding. Poverty in the sense discussed here or relative incomes measured using PPP calculations may not, however, be the best indication of need for any specific form of special treatment, so the 'right' answer may be that no measure of per capita income should be used for these purposes. Depending on the purpose, what is needed may be direct measures of health, education, access to services, or level of agricultural or industrial development. As Case Study 2, by Donald Roy, shows, these will require different price comparisons.

How does spending on public services differ between countries?

Making sectoral comparisons between countries raises similar issues of finding ways to measure spending on specific goods or services, not average incomes. The difficulties in answering the very political issue in the UK of whether it spends too little on health relative to other developed countries are illustrated in Case Study 2. This demonstrates the variations in results from using different measures, and in particular the risks of using aggregate measures, which in this case exaggerated significantly the measured difference between the UK and other developed countries.

Implications

Most users of comparison measures cannot design their own index and collect new data. But users must take responsibility for understanding and specifying exactly what they are trying to measure, then choosing the best available measure for that purpose. They must give clear information about how the measure differs from what would be most appropriate and how this is likely to bias the results; and, if possible, they should try alternatives to check how sensitive the results are. They should also recognise those questions that cannot be answered by income comparisons (however defined).

Those who compile data must specify the assumptions and methods sufficiently clearly for users to be able to use them appropriately.

The differences in countries' spending patterns (and in the choices made by their statisticians) are normally believed to make comparisons across time (have incomes increased?) more reliable than across space (which countries are now rich or poor?). But improved analysis leading to the compilation of more or different data (e.g. more understanding of different spending patterns), may limit our ability to make comparisons across time, even though it may improve comparisons within or across countries.

The fact that perfect measurement is unattainable does not mean that nothing can be said. Some results are robust despite all these qualifications: countries now considered to be middle- or higher-income developing countries have passed, on all measures, the levels of countries considered to be developed in the 1960s when measurement and assistance for development began. Some of the countries at the lowest levels on all the measures at the beginning of the process of measurement and policy remain there to this day.

By Samuel Moon, ODI Research Officer (s.moon@odi.org.uk), Sheila Page, ODI Senior Research Associate (s.page@odi.org.uk); Barry Rodin, Chief Economist, ECA International, and Donald Roy. For more information, contact s.moon@odi.org.uk.

References

Appleton, Simon; Emwanu, Tom; Kagugube, Johnson and Muwonge, James (1999) *Changes in poverty in Uganda, 1992-1997*. Centre for the Study of African Economies Working Paper Series, Working Paper 106. Oxford: Centre for the Study of African Economies, University of Oxford (www.bepress.com/csae/paper106).

Asian Development Bank (2008) 'Research Study on Poverty-Specific Purchasing Power Parities for Selected Countries in Asia and the Pacific'. Manila: Asian Development Bank.

Atkinson, Anthony B. and Brandolini, Andrea (2009) *On Analysing the World Distribution of Income*. Bank of Italy Working Paper. Rome: Bank of Italy, Economic Research Department.

Chen, Shaohua and Ravallion, Martin (2008) *The Developing World is Poorer than we thought, but no less successful in the fight against poverty*. World Bank Policy Research Working Paper 4703. Washington, DC: World Bank.

Deaton, Angus (2000) *Counting the world's poor: problems and possible solutions*. Princeton University RPDS Discussion Paper No. 198. Princeton, NJ: Princeton University.

Ferreira, Francisco H. G. and Ravallion, Martin (2008) *Global Poverty and Inequality*. World Bank Policy Research Working Paper 4623. Washington, DC: World Bank.

Levy, Stephanie (2009) 'International Poverty Line: What's New in Counting the Poor?'. Paper prepared for IPIC seminar.

Moon, Samuel (2007, mimeo) 'Poverty Line Methodology Issues'. Kampala: Ministry of Finance, Planning and Economic Development.

ODI (1988) 'The Rich and the Poor: changes in incomes of developing countries since 1960'. ODI Briefing Paper. London: Overseas Development Institute.

Ravallion, Martin (2009) *A Comparative Perspective on Poverty Reduction in Brazil, China, and India*. World Bank Policy Research Working Paper 5080. Washington, DC: World Bank.

Ravallion, Martin, de Hoyos, Rafael E. and Medvedev, Denis (2008) *Is the Developing World Catching up?*. World Bank Policy Research Working Paper 4733. Washington, DC: World Bank.

Roy, A. D (1982) 'Labour Productivity in 1980: an international comparison', *National Institute Economic Review* No. 101, August.

Wanless, Derek (2002) 'Securing our Future Health: Taking a Long Term View'. Interim Report.

World Bank (2008) ICP data for 2005 round (www.worldbank.org/data/icp).

World Bank, International Comparisons Programme, http://go.worldbank.org/MP1FBPYUJo