Use of Penn World Tables for International Comparisons of Poverty Levels: Potential and Limitations

Bettina Aten and Alan Heston[1]

Introduction

Individual scholars and expert groups have considered both physical and monetary measures to define a poverty line. Physical measures might be based upon caloric intake, or as considered in India, the number of square meals (thoughtfully defined) in a sample period and the purchase of clothing, with some experts also arguing for an inventory of physical household assets. The attraction of choosing physical measures is that they appear to avoid the necessity of converting currencies across countries to a common measure. However, the survey problems of identifying those below such thresholds, or of choosing equivalent food baskets across space to obtain the same caloric intake, involve problems as knotty as those using monetary measures. In any event the focus of this paper is on monetary measures of poverty and takes up three interrelated issues of comparing poverty levels within and between countries.

The first area discussed in Section A is the set of expenditure weights used to generate the aggregate purchasing power parities (PPP) used for converting currencies to a common measure. In Section B of the paper the particular approach to estimating PPPs in the Penn World Table (PWT) is described, including a discussion of the benchmark estimates, reconciliation methods, short-cut estimates and comparisons with other data sets. In Section C the nature of the item prices that enter into any calculation of PPPs, including PPPs for sections of the population (e.g., the poor, rural residents) is discussed with some illustrations.

A. Expenditure Weights used to Aggregate Heading Parities

The factor used to make currencies comparable across countries, whether it be an exchange rate or purchasing power parity, does not change the underlying economic inequalities around the world. However, it does significantly affect the perception of economic size and of poverty. At PPPs the economies of China and India are ranked among the world's top seven, but not so at exchange rates. Similarly if one took a simple poverty line like one quarter the world per capita average, then at exchange rates in 1980, the proportion in poverty would approach 80% in Asia and other poorer continents, while at PPPs it would be closer to 50%.[2] So use of a particular conversion factor can significantly change our perception of the extent of the problem.

---

[1] Aten teaches at Bridgewater State College and is on leave doing research at the Bureau of Economic Analysis, US Department of Commerce; Heston teaches at the University of Pennsylvania. Support from the National Science Foundation is gratefully acknowledged. Parts of this paper draw on a presentation to the World Bank, ABCDE conference in May, 2001.

[2] These comments and parts of this section are based on Heston (1986).

Use of the PPP for consumption is clearly more appropriate than say the PPP for gross domestic product (GDP) for converting any common international poverty line into local currencies. Both are available in PWT. Are there other PPP concepts that should be considered? PWT and the benchmark International Comparison Program (ICP) estimates are based on plutocratic weights. An obvious alternative would be to use consumption weights of the poor. This is a fairly natural extension of what is done in temporal price indexes where one might estimate a consumer price index (CPI) for subgroups of the population like the elderly, or use democratic versus the more common plutocratic weights. Heston (1986) experimented with this approach using expenditure weights of the poor (approximately the lowest quintile) to estimate what could roughly be called a PPP of the poor. It is reported below.

Consumption PPPs with Weights of the Poor

As will be clear from the discussion in Part B, this experiment is only possible for ICP benchmark countries where parities for detailed headings of consumption are available. The exercise reported below was based upon 55 of the 60 countries in the 1980 ICP benchmark (Eastern Europe was excluded). Two limitations of this exercise should be mentioned. First, the exercise was not on input parities at the basic heading level of 100 or so categories of consumption, but on parities at the level of major expenditure components, about 10-20 summary categories. These summary category parities for 1980 had been estimated by the Geary-Khamis (G-K) aggregation method. It is possible that the over-all quantitative effect of alternative weights would differ from those using detailed parities, but it is not likely that the use of G-K parities would change the qualitative findings.

The second limitation of this exercise is that it used representative expenditure weights of a very diverse character. For Africa, the expenditure distribution was of estate workers in Malawi. In Asia, the 3rd decile of rural workers in India; for South America, the lowest quintile in Brazil; and in high income countries, the 1960 expenditure distribution for those in poverty in the United States was used. The overall results are given in Table 1.

Table 1: Ratio of PPPs of the Poor to those of the Average PPPs, 1980

Region	Number of 1980 Countries	Ratio (Poor/Average PPP)
Africa	14	.876
Asia	5	1.205
South America	17	1.310
Developed Countries	19	1.064

The major surprises were that the weights made as much difference as indicated in Table 1, and that the results were so different in Africa. What drives the results in the developing countries is the relative price of food, which is typically high compared to other headings of consumption. This means that for Asia and South America, using a consumption PPP based upon an expenditure distribution closer to the poverty level would increase the number in poverty.

If we believed the African figures they would suggest the number in poverty in 1980 was substantially overstated, especially compared to Asia and Latin America. Prices of many food items were controlled in Africa in 1980, and while efforts were made to obtain prices of food as a weighted average of ration and free market prices, it is doubtful that this was actually done in many countries. The average ratios in Table 1 cover up a fair amount of country variance, but there were only 2 of the 14 countries in Africa where the parity of the poor was higher than of the national average. However, it seems probable that the African PPP for the poor is low because ration prices entered into the estimation with more weight than justified by their quantitative importance. A study carried out by Yonas Biru and Sultan Ahmad at the World Bank, based on prices collected in a number of African countries for 1985, produced a result similar to that reported above for Asia and South America .

There is little basis for strongly defending the numerical results of this exercise. However, there seems to be a larger quantitative effect than might be expected when we use weights closer to expenditure patterns of those in poverty. This seems to be an area worth more exploration.

Which Consumption Concept Should be Used?

The benchmark framework adopted in 1968 used an ICP concept of consumption, namely private consumption expenditure plus those parts of government expenditure on health and education that accrued directly to households. At the time this was not the practice in the 1964 SNA, but was recommended by the Eastern European countries participating in the 1970 benchmark comparison. The reason was that the ICP approach would allow more meaningful comparisons of consumption and its components across countries that financed their education primarily from public funds, versus countries that mainly used private expenditures. In the early benchmark comparisons both the ICP and SNA consumption concepts were presented.

Unfortunately, PWT versions 3 to 5.6 had to be based on the SNA concept of consumption, namely final household consumption; this was because the detail necessary to provide both concepts was not available for years and countries other than benchmark ones. Since the PWT treatment corresponds to the basis for most poverty estimates, it means that the consumption PPP from PWT is generally the best conversion factor for international poverty accounts available from that data set.

In 1993 the SNA adopted the original ICP convention and distinguished between Household Final Consumption Expenditure, the old SNA concept, and Household Actual Final Consumption, which differs primarily in including the expenditures of government and non-profit institutions on health and education that directly accrue to households. The OECD countries have all adopted the 1993 SNA, and PWT6.1 (2002) now suffers in that it provides Household Actual Final Consumption for them, while using Household Final Consumption Expenditure, by necessity, for other countries. This is not a major problem for international comparisons of poverty because the poorest countries are primarily non-OECD countries, but clearly presenting both consumption concepts in PWT should be the goal.

B. A Brief Description of PWT Methodolgy

This section sets out the main differences between PWT and other data sets that might be used for studies of international poverty. We first distinguish between the treatment of benchmark and non-benchmark countries. The method of producing current and constant international price estimates is treated next along with the principal differences between the PPP estimates of the World Bank and PWT. A more detailed version of the materials described in this section is provided in the documentation of PWT 6.1[3] so this discussion will be brief.

Benchmark and Non-Benchmark Countries

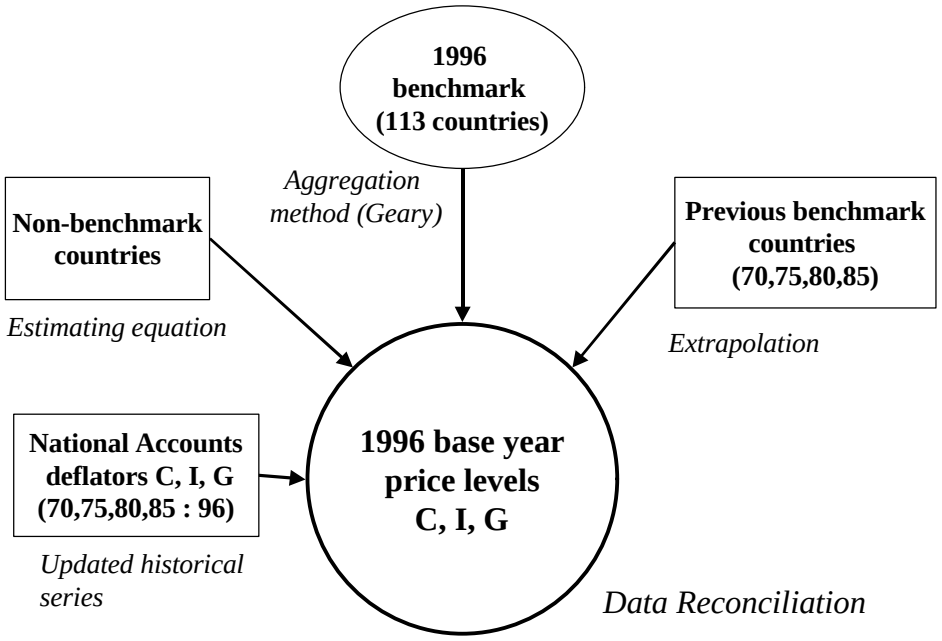
Benchmark ICP comparisons have been carried out for over 100 countries, some for just one year, and some for as many as eight years since 1970, originally at five-year intervals, and now every three years for the OECD countries.[4] Benchmark comparisons typically involve detailed price comparisons representing 150 or more basic headings of expenditure on consumption, capital formation and government. Beginning in 1980 these benchmark comparisons have been organized regionally with various procedures built into the process so that links could be established between countries in different world areas. Some links were provided by countries in both OECD and other groupings, as for example Austria with countries of Eastern Europe, and Japan with the Economic and Social Commission for Asia and the Pacific (ESCAP).

Unfortunately, the last ICP benchmark that represented most of the world regions for a particular date was 1985; it was incorporated in PWT 5.6, with later regional benchmark data. For PWT 6.1, a world comparison was cobbled by using 1996 OECD estimates for member countries plus an equal number of formerly planned economies. Several Latin American countries also made estimates for 1996, and it was possible to update 1993 estimates for the ESCAP countries, Africa, the Middle East and the Caribbean to 1996, for a total of 113 countries, albeit at the level of only 36 headings of expenditure[5].

Figure 1 illustrates the inputs and procedures used in PWT to obtain the initial 1996 base year price levels (or PPPs) for the three components of GDP: consumption, investment and government.

[3] Technical Documentation, About PWT, in http://pwt.econ.upenn.edu.
[4] Actually the European Union countries have been carrying out annual estimates since 1993 where about one third of the underlying items are priced each year, and the remainder updated from the previous years by appropriate time-to-time indexes.
[5] In 1985 there were only 64 countries but a total of 139 basic expenditure headings.

Figure 1. Base Year PPPs



First, the ICP benchmark data for the 113 countries are aggregated to the level of C, I, and G using the Geary-Khamis (G-K) method and weights, termed super-country weights, that assign proportional representation of the benchmark countries relative to the world. The World Bank has used a different aggregation method and a different weighting scheme, one that assigns equal weight to each country over all of GDP, so that small countries such as Belize and Luxembourg will have the same importance over all headings as larger countries such as Mexico and Germany. The use of super-country weights in the G-K system provides continuity with previous versions of PWT.

The second step is to estimate the PPP of C, I and G for the non-benchmark countries.[6] In recent versions of PWT these estimates have been made in two stages. First, an estimate of the PPP for Domestic Absorption is made based upon the relationship between various cost of living measures and the PPP for GDP for benchmark countries. The values of these post adjustment indexes for the non-benchmark countries are then used in the estimating equation to obtain their Domestic Absorption. This may be contrasted with the method used by the World Bank, also a short-cut approach, but one that uses an equation involving education and nominal

[6] There are also countries for which benchmark results are not available but some studies have been made, notably China and Taiwan. For details see PWT6.1.

income but no direct information on prices in non-benchmark countries. In addition, the World Bank does not make estimates for C, I, and G, only for GDP, whereas in PWT the component PPPs are estimated again using a relationship derived from the benchmark countries.

The third and final step is to collate the 1996 benchmark PPPs, the non-benchmark PPPs and the PPPs from previous benchmark countries that may or may not be part of the 1996 ICP. When countries have multiple benchmarks, the relative PPPs of two countries in two benchmarks usually differs from what would be predicted from relative price movements in the two countries. For example, if the GDP deflator in country A rises by 20% between two benchmarks and that of B by 30%, then one would expect the $PPP_{B/A}$ to rise between two benchmarks by about 8.3% [(1 - 1.3/1.2)*100]. In fact the two estimates will differ, often by 5 to 15% or more in either direction.

To deal with this empirical finding we use a reconciliation process[7]. The basic idea is to bring previous benchmark estimates of PPPs to a common year by use of the national accounts deflators. For countries with several benchmarks it is necessary to average the different PPP estimates and this is done by giving more recent estimates somewhat greater weight. The reconciled past and present benchmark PPPs, together with the non-benchmark short-cut PPP estimates, and the national accounts expenditure data, become the inputs to another multilateral aggregation procedure (G-K method, super-country weighting) that will generate the GDP PPPs and international dollar estimates for C, I, and G for the 168 countries in 1996.

It should be noted that these estimates will not necessarily correspond to the initial benchmark comparison for 1996 because both non-benchmark and previous benchmark countries are now included. The World Bank does not attempt this reconciliation process.

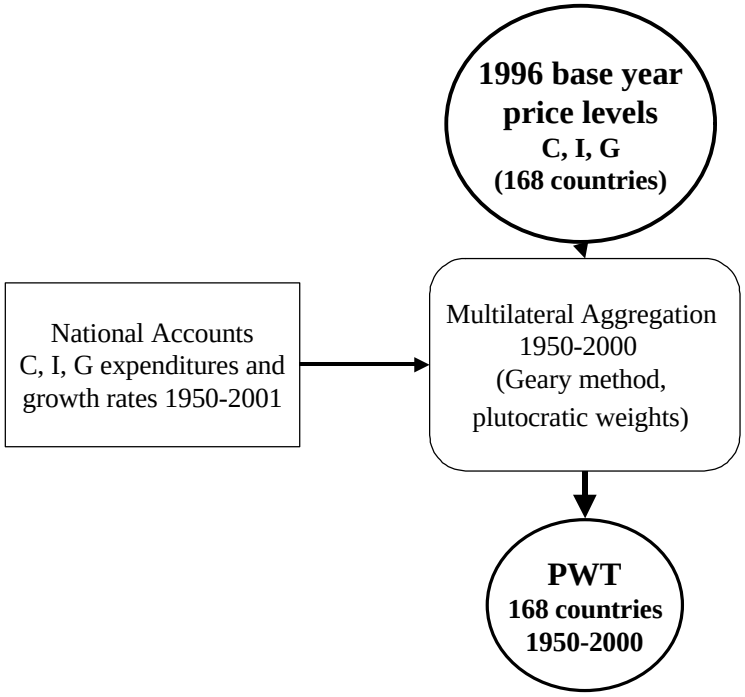
PWT Estimates in Other Years

Frequently, international comparisons of poverty, and of wealth, are made at different points in time. One advantage of PWT as a data source for the PPP for such estimates is that it provides a continuous series from which erratic movements that may occur using benchmark estimates in two different years have in effect been removed[8]. Figure 2 illustrates the procedure to obtain the current and constant price series in PWT over time.

[7] This reconciliation process was called 'consistentization' in previous versions of PWT, but Robert Summers has reluctantly given up the term.

[8] The reconciliation process does not remove erratic movements that originate in the national accounts series.

Figure 2. PPP and GDP Estimates Other Years



For 1996, we have the set of 168 benchmark, non-benchmark and previous benchmark countries and their component PPPs. For other years, we move the 1996 PPPs backwards and forwards by the changes in the national accounts deflators for each component of each country relative to changes in the United States. These become the input PPPs that, combined with the current price national accounts of each country, permit a new multilateral aggregation (G-K method, super-country weights) for each year. The result is a set of GDP PPPs and international price estimates of C, I and G for the 168 countries for 1950-2000.

Several different constant price measures are provided in PWT. It is not clear that researchers would want to use these in poverty comparisons, so the following discussion is highly condensed. A Laspeyeres type measure is given that takes the real value of the components in each year and moves them backward and forward by the national accounts growth rates of the components. The resulting estimates are summed with the net foreign balance in 1996 prices to obtain the GDP in each year. Because the weights of C, I and G in international prices will not necessarily be the same as those in national prices, the growth rate of GDP In PWT will not be identical to that in national prices. In this PWT differs from most other series and this should be understood in research making use of the growth rates implicit in PWT. The same is true of the chain

index in PWT. The chain index applies the national accounts growth rates to the component shares in international prices, derived from the current year multilateral aggregation, obtaining a growth rate for Domestic Absorption (DA) for each pair of consecutive years.

The main differences between PWT and World Bank PPPs can be summarized as follows:

1. The initial *aggregation method* or price index number formula that is applied to the benchmark countries is not the same: PWT uses the G-K aggregation with plutocratic weights.
2. Estimates for *non-benchmark* countries are made using short-cut methods, but the equations and variables differ: the World Bank uses education and nominal incomes whereas PWT uses information on prices and no education variable.
3. Information on *previous benchmarks* is not used in the World Bank, but is collated and reconciled in PWT.
4. The *current price series*: PWT estimates PPPs and international prices for each component in each year, whereas the World Bank obtains the 1996 GDP PPPs and applies national accounts growth rates to obtain other years.
5. The *constant price series*: PWT's Laspeyres series is based on the growth rate of C, I and G from the national accounts plus the net foreign balance, the World Bank uses GDP growth rates.
6. *Chain series*: PWT provides a chained constant price series using component shares in international prices for each year.
7. *Consumption PPPs*: PWT provides the PPP and the constant and current international prices for consumption as well as for GDP for all countries and for as many years as there are national accounts series available.

C. What Prices are Used?

Estimation of basic heading or aggregate PPPs is not easy, even if there were agreement on all the methods to be employed. Comparisons of prices of comparable goods and services below the detailed heading level are the basis for estimating benchmark PPPs. However, the devil is in the details. And this is important for the benchmark ICP comparisons, and consequently for the quality of PWT estimates of consumption PPPs, and our ability to compare poverty levels in common currency units across countries. One of the key issues involves the common practice of using national average item prices from each country. This sounds sensible enough. But when we are focused on a particular population group, namely those in poverty, do national average prices make sense? Probably not, and that is a problem, along with selection of items and outlets, that we take up in this section.

Items Consumed by the Poor

Consider an item about which millions of us are experts, haircuts. On a summer day in Beijing one can get a haircut at shops with varying amounts of amenities. Excluding hotels the charges might range from 10 to 15 yuan in shops down to 1 or 2 yuan for no-overhead service on the street. Most PPP estimates will choose a shop that

might be found in a range of other countries. As long as the prices in such shops represent the relative costs of these services in different countries, the comparisons may still be reasonable, even if they do not explicitly price street barbers.

However, the existence of phenomena like street barbers raises some disturbing issues. If we consider provision of a minimum bundle of necessary goods and services as the basis for determining the cost of a poverty bundle, then what do we do about items like street haircuts that may only be consumed by the poor? Other examples include rice with broken grains that is indifferently sorted and cleaned, inferior grains like *ragi*, in India, second hand clothing or cloth remnants, and a wide variety of inferior or makeshift housing. The present state of PPP estimation, at best, only represents consumption of these items by goods and services in the relevant expenditure heading that are available in a wide range of countries.

The Prices to the Poor Question

Outlets and items used in the ICP are those thought important in the expenditures of each country, and often the items representing a heading in one country will be different from those in another country. But do the poor pay different prices for the same items than the middle and upper classes? Those in the PPP estimation business are silent on this issue, and if asked, will say we would be glad to have such data if only countries collected them.

In an early study, Kunreuther (1973) set forth a simple model to examine this question that took into account size of packaging, type of outlet, and inventory costs of large package sizes or bulk purchases. He found that in New Haven the same package size was more expensive in small stores than chain stores and of course that price per physical unit declined with increasing package size.[9] The link to poverty occurs in where stores are located, where the poor make purchases and the size of package they purchase. His result was quite clear. The poor purchased in smaller size packages in smaller stores. Why? Chain stores were not in poor neighborhoods and the poor had less access to their own transport to travel to larger stores. The poor interviewed in the Kunreuther study traveled smaller distances than the more affluent, and had less ability to store goods. In addition to the storage constraint, the poor had weekly per capita purchases that were about 2/3 those of the middle class sample interviewed. The poor also made more frequent purchases suggesting that storage and liquidity constraints may have both operated to produce purchases of smaller size packages.

That was New Haven in 1973. A study in northeast Brazil by Musgrove and Galindo (1988) reported a somewhat different result when they looked at small and large stores in large, medium and small cities. Their study was in 1985 and in an attempt to overcome the effects of the overall rapid inflation in Brazil in those years they concentrated the survey into two weeks in a month with only (!) a 5.41% price increase. Whereas Kunreuther found that neighborhood stores sold the same size package at a

[9] In Kunreuther's sample the price per unit over the range of package sizes was from 50% to 75% going from largest to smallest size. The sample of poor and middle class respondents were well aware of the range of sizes available and the differences in price per physical unit even though this was before mandatory displays of this information in chain stores. In the sample of neighborhood stores, about 20 to 40% stocked the largest package size for each of the 8 items sampled by Kunreuther.

price typically 10 to 15% higher than the chains, Musgrove and Galindo did not find such a consistent pattern, with some items like manioc flour, being sold at lower prices by small retailers. Further, they report that for items sold in bulk, like beans or rice, the price per unit was the same whether the size was a cup or a much larger quantity. A limitation of their study was that it relied solely on the response of storeowners. They did not have direct information on where the poor made their purchases or at what prices. However, it does appear that in urban areas of Brazil the poor do not face different prices for the same goods as the rich, in part because many neighborhoods have a wide range of socio-economic groups living in close proximity, which is not to say that lack of capital does not constrain the size of unit that the poor purchase.

V. Rao (2000) dealt with this question in a study of villages in South India and found that, because the poor buy in very small quantities, the price paid per kilogram of basic food items is higher than for the middle classes. For example, a kilo of yellow split peas would cost Rs. 28, and a 100 gram purchase, Rs. 3.50. While poor families in a week may buy grains in sufficient bulk, important commodities like *pulses* may, as in the above example, have a 20% higher unit cost. Similarly cooking oil is often purchased by the poor in 100 gram lots, raising the unit price. In rural areas, there may be little effect of outlets, but Rao found a significant effect of size of purchases on the cost of a given quantity of consumption goods between the very poor and better off villagers.

What of urban areas of India? Anecdotal evidence abounds. Sales of individual cigarettes at small street stalls reveal the same higher costs per unit as a correlate of low income and/or little liquidity.[10] Even when prices per kg are similar for larger and smaller size purchases, there is typically in India a valuable gift with the large package.[11] Is there an outlet effect for the same size of purchase such as in New Haven, but which was not systematically evident in northeast Brazil? Certainly the ICP framework has in the past provided no basis for examining this issue.

Even if there is no outlet effect, if the poor pay as much as 10-20% higher prices because they buy in smaller quantities, this would be useful to know. Would it make any difference in measuring the number in poverty? It would clearly affect the total count and it would also affect a comparison of countries if these effects are not similar in magnitude across countries. The studies cited above suggest the size of purchase and outlet effects may be different across countries, in which case it would be even more desirable to have country research on these issues.

Continuing this theme we report some results from an exercise carried out by Michael Perling, an undergraduate having low budget travel time in Asia and need for course credit for an independent study. With limited resources but an enthusiastic traveling companion, he collected some 2787 price observations on13 commodities and services in rural and urban areas of China, Hong Kong, Thailand, Malaysia and Singapore in the Spring of 2002. While necessarily anecdotal, these results are

[10] In fact the pricing of an individual cigarette at Rs. 2 can be fairly close to a package of 10, that may cost something over Rs. 15. And *panwallas* may use a low price of a single cigarette as a loss leader. But informants in Brazil and Egypt suggest that the mark-up or the single cigarette is typically 10-20% above buying a package.

[11] For example a 1 kilogram package of cooking oil may sell for Rs. 55 and a 5 kilogram package at Rs. 225, but the latter will include a plastic bucket valued at Rs. 90 by the seller, but at cost perhaps Rs.50. A significant percentage of larger size consumer items in India are discounted in this tied manner.

suggestive of the type of survey that would address some of the issues raised above. Before highlighting a few of his findings, a slight digression is in order on the role that hedonics can play in estimation of parities for ICP purposes as well as for prices for particular population groups.

The Bureau of Labor Statistics (BLS) Approach to Spatial Price Comparisons

ICP type price research looked for models in terms of known frameworks for price collection for country consumer price indexes (CPI). CPI methodology typically either averages prices across outlets in a city and then takes the time to time price relative for the item, or takes a price relative at each outlet and averages the relatives across outlets in a city. In either case, information that might have been available on outlet type and average quantity purchased is discarded in the aggregation process. However, in general, price collectors know the location of their outlets, and could easily learn about typical sizes of purchase for items where it is relevant.

The BLS in the United States changed its framework for CPI price collection in a way that at first glance made the problem of using their data to compare prices across space very difficult. There is a sampling frame at which the price collector checks off for each entry level item (ELI) the outlet, the size, the type of package and other information about the volume seller within the ELI as indicated by an outlet employee. When this framework was adopted by BLS in the 1970s it seemed not to lend itself to place to place comparisons because collectors were not asked to price the same item in different outlets. There is no way of knowing in this framework whether, for example, the type of soft drink priced in supermarkets in Denver is the same as those priced in Chicago.

But Kokoski, Cardiff and Moulton (1994) and Kokoski, Moulton and Zieschang (1999) demonstrated that the framework of the CPI lends itself clearly to a hedonic approach. In fact, Kokoski *et al*, began experimenting with the hedonic approach that was also part of early ICP work, namely the Country Product Dummy method (CPD) developed by Robert Summers (1973). The version that Summers used was a very straightforward hedonic regression model akin to those used for temporal studies (Griliches, 1990, Triplett, 1990, Berndt, 1995). The prices are regressed against the two sets of dummy-variables as given in equation (1) below: one set contains a dummy variable, D_j for each country other than the numeraire country (country 1), and the second set with a dummy for each item specification, z_i.

$$(1)\ \ln p_{ij} = \sum_{i=1}^{n} b_i z_i + \sum_{j=2}^{m} a_j D_j + e_{ij}$$

The transitive price parities, a_js, are the logarithms of the estimated country parity for the heading relative to the numeraire country. The item coefficients, the $\hat{a}_i$s, are the logarithms of the estimates of the average item price in the currency of the numeraire country (which could also be a regional currency).

The innovation of Kokoski and colleagues was to apply this data to estimating internal price parities by BLS city using the entry level item (ELI) characteristics of the prices being collected. The basic idea was similar to the CPD procedure. We may not be able to match the specific apples priced in Philadelphia with those priced in Los Angeles. But across all the BLS cities, so long as there is an overlap of specific apples priced in some cities, then a parity can be obtained for all apples between any pair of cities. If Apples is the basic heading, and price per kilo is the unit, then for each price observation in the ELI (Apples), there would be a code for outlet type, city, and item specification (Fuji, Rome, Granny Smith, Delicious, McIntosh, etc.)

The application of this hedonic framework that is proposed for a poverty PPP is set out in (2) below. The subscript *j* may refer to countries as in the CPD method or as in the BLS formulation, *j* may refer to regions within a country.

$$(2)\ \ln p_{ikj} = \sum_{i=1}^{I} b_i z_i + \sum_{k=1}^{K} b_k z_k + \sum_{j=2}^{m} a_j D_j + e_{kj}$$

The subscript *i* refers to the outlet type (e.g., low, average or high income outlet), while the *k* are item specifications. With this information a simple hedonic regression could tell us whether coefficients for dummy variables in outlets in poor neighborhoods were significantly higher than in middle class neighborhoods for different types of items.

Illustrations of Subnational PPPs by Population Group and Geography

Perling (2003) estimated a version of equation (2) for each of the thirteen items for which price information was collected. A few of his findings illustrate the way in which check-list type price information may improve our understanding of how prices may differ across socio-economic groups and regions[12]. Table 2 provides summary results for three items in Perling's study, a durable item, batteries, a perishable, onions, and a service item, haircuts for men. In Table 2 the price level of each item is presented. The base for the comparison prices collected in outlets in a middle class area of Chengdu city in China. For Bangkok, the PPP of the Thai Bhat to the Chinese Yuan is divided by the exchange rate and expressed as a percent. For example, the entry of 227 for a kilo of onions in Bangkok means that it costs the Bangkok middle class 2.27 as much as the price for a kilo of onions in Chengdu at exchange rates.

[12] In the next set of benchmark comparisons for 2004 the term "structured product description" (SPD) will be used in place of item specification. Item specifications in the ICP have often been highly specific, and SPDs will introduce more flexibility into price collection by listing variations that are acceptable. In this way an SPD is somewhere between the ELI and the traditional item specification. It may be noted that the Fundação Instituto de Pesquisas Economicas (FIPE) regularly estimates a consumer price index for São Paulo, using SPDs that permit the type of analysis described here. Professor Heron Carmo, Director of the CPI for FIPE kindly provided some sample data for the 55 districts of the city covering a range of outlets, brands and varieties of goods and services. Experiments have been carried out for a number items, including chicken, milk, dental services and shampoo and were quite promising and are continuing.

Table 2. Price Levels in Selected Asian Markets, Spring 2002
(Base Chengdu, Middle Class = 100)

	Batteries	Onions	Haircuts
Bangkok – Middle	79	227	208
Bangkok – Poor	79	306	102
Singapore – Middle	135	500	582
Singapore – Poor	101	312	406
Shanghai - Middle	110	151	160
Shanghai – Poor	84	135	107
Fuli – Rural	59	85	22
Shenzen – Middle	109	164	126
Hong Kong – Middle	118	477	835
Chengdu – Middle	100	100	100

In Table 2, differences within a country can be interpreted as involving mainly a price difference for batteries (brand and outlet type is held constant) and onions (type of outlet held constant), and both price and quality for haircuts. The only rural setting is Fuli in China, which represents the lowest price level for all items. In the case of batteries, this may represent quality differences that are not observed (shelf life date, for example). The difference between poor and middle class in the sample represents judgments about the neighborhoods where the outlets were sampled. For a service item like haircuts, is it sensible to compare middle class neighborhoods across China and across countries? The results are not implausible.

Taken at face value, what would the information in Table 2 on China suggest about the geographical distribution of poverty within a country? As is well known, the poverty count is inversely related to the level of income in states, provinces or any other sub-national unit. This fact is often used to justify policies that promote overall economic growth in a country as the most useful way to reduce poverty. However, it is not inconsistent with that position to also try to measure the poverty in different regions better than we do. In some cases better measures of the geographic dispersion of poverty may facilitate targeted policies that can supplement income growth in raising the economic, educational, or health status of specific groups. (See for example, Bigman and Fofack, 2000) The numbers for China in Table 2 suggest that taking account of price differences within China would reduce the poverty count in rural areas and poorer cities and raise them in better off cities.

A study of the poverty line in the United States by Aten (1996) illustrates another aspect of the regional problem. The notion of a national poverty line has been under review in the United States for a number of years but there remains a lack of consensus on exactly what to do. In the meantime, a number of poor in regions like the Dakotas are over-counted and those in large cities undercounted. In her work Aten using the Kokoski, Cardiff and Moulton (1994) study of inter-area price differences in the United States calculated the cost of the national poverty bundle in1987 which was then $5778 per person. This bundle cost $4867 in the North-Central region versus just over $6970 in San Francisco and the New York SMA, hence the likely over-count of those in the North-Central and South Regions compared to most large U.S. cities. While the government may have had political reasons to shy away from sub-national poverty lines,

it has certainly not stopped a number of private firms from selling their estimates of how costly it is to live in different parts of the United States.

One conclusion is that it would be desirable to build up price levels by geographical region and population groups like the poor, perhaps by estimation of hedonic regressions for a number of goods and services. These hedonic equations would explain price by size of package and a series of dummy variables including type of outlet, region of the country, rural versus urban location, and within urban areas, poor and other neighborhoods. If the relative importance of different size purchases in rich and poor neighborhoods is known, it could be used to sharpen the PPP estimates for the poor. This would permit regional estimates of price levels, real income and numbers in poverty. National totals of poverty would then be summed up from the regional numbers. This would serve national statistical goals and provide at least as good a basis for policy decisions as obtained using a national poverty line and then estimating the numbers in poverty in each region.

Concluding Remarks

This paper has reviewed some of the strengths and limitations of PWT for providing suitable PPPs for international poverty comparisons, described some aspects of PPP estimation compared to alternatives and suggested several ways in which estimation of PPPs for the poor might be improved in benchmark ICP comparisons. Improvements can be summarized as a) to use expenditure weights of those in poverty in obtaining an aggregate PPP for consumption; and b) to take into account prices paid by the poor. If direct surveys are not available, an indirect approach is to use existing price surveys to identify the location, outlet and type of neighborhood where prices were collected, and the typical size of purchase. Research using an approach like this may permit us to estimate PPPs for private consumption that are more appropriate for the poor, both within and between countries. The first suggestion can be carried out for benchmarks but not other countries in PWT. The second suggestion would try to improve the underlying price data entering into PPP and subsequently PWT calculations. At present, PWT consumption PPPs provide a basis for conversion of international poverty lines into national currencies that are fairly stable over time. However, their underlying price base is no stronger than existing benchmark comparisons for which happily a major international effort is underway for their improvement.

References

Aten, Bettina (1986), “Some Poverty Lines are More Equal Than Others,” Presented to the Western Regional Science Association, February. Discussion Paper of the Center for International Comparisons at the University of Pennsylvania, CICUP 95-5.

Berndt, E.R., Grilches, Z., and Rappaport N.J. (1995), Econometric estimates of price indices for personal computers in the 1990s, *Journal of Econometrics*, 68, 243-68.

Griliches, Z., (1990). 'Hedonic Price Indices and the Measurement of Capital and Productivity: Some Historical Reflections'. In *Fifty Years of Economic Measurement: The Jubilee Conference of Research in Income and Wealth*, *Studies in Income and Wealth*, Vol.54, NBER (eds. E.R. Berndt and J.E. Triplett), pp.185-206. Chicago University of Chicago Press.

Heston, Alan (1986), "Some Problems of Comparing Poverty Levels Across Countries," paper presented at Seminar on *Use of Purchasing Power Parities*, Luxembourg, Eurostat, November.

_____, Robert Summers and Bettina Aten (2001)**,** "Stability of Country Price Structures: Implications for Spatial-Temporal Comparisons," UN Economic Commission for Europe Statistical Reporter, May, 2001. Also on OECD website.

Kokosky, M, P. Cardiff and B. Moulton (1994), "Interarea Price Indices for Consumer Goods and Services: An Hedonic Approach Using CPI Data, BLS Working Paper 256.

Kokosky, M., B. Moulton and D. Zieschang (1999), "Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Aggregation," in International and Interarea Comparisons of Income, Output and Prices, A. Heston and R. Lipsey editors, NBER, Studies in Income and Wealth Volume 61, U of Chicago Press.

Kravis, I.B., A. Heston and R. Summers (1982) World Product and Income, World Bank and Johns Hopkins University Press, Baltimore.

Kunreuther, Howard (1973), "Why the Poor May Pay More for Food: Theoretical and Empirical Evidence," The Journal of Business.

Rao, Vijendra (2000), "Price Heterogeneity and 'Real' Inequality: A Case-Study of Prices and Poverty in Rural South India," Review of Income and Wealth, June.

Summers, Robert (1973), "International Comparisons with Incomplete Data," Review of Income and Wealth, March.

Triplett, J.E. (1990). 'Hedonic Methods in a Statistical Agency Environments: An Intellectual Biopsy'. In *Fifty Years of Economic Measurement: The Jubilee Conference on Research in Income and Wealth, Studies in Income and Wealth*, Vol 56, NBER (eds. E.R. Berndt and J.E Triplett), pp.207-238. Chicago: University of Chicago Press.