the review of income and wealth

Review of Income and Wealth
Series 61, Number 3, September 2015
DOI: 10.1111/roiw.12116

IMPROVING INTERNATIONAL COMPARISONS OF PRICES AT BASIC HEADING LEVEL: AN APPLICATION TO THE ASIA-PACIFIC REGION

BY ROBERT J. HILL*

University of Graz

AND

IQBAL A. SYED

University of New South Wales

The International Comparisons Program (ICP) run by the World Bank compares prices and real incomes across countries, and plays a pivotal role in the Penn World Table. Using a unique dataset consisting of over 600,000 price quotes from nine countries in the Asia-Pacific region, we consider ways of improving the basic heading price indexes that form the building blocks of ICP. Current ICP methodology computes these price indexes using the country–product–dummy (CPD) method applied to the country average prices. We contrast this approach with: (i) a weighted version of CPD; (ii) CPD applied directly to the individual price quotes; and (iii) extended versions of CPD that include adjustments for unrepresentative products, urban–rural price differences, and different outlet-types. Also considered are new CPD-based methods for measuring urban–rural price differences, and the implications of our findings for the downward revision in China's GDP in ICP 2005.

JEL Codes: C43, E31, O18

Keywords: China/Chinese GDP, country-product-dummy method, International Comparisons Program, representative and unrepresentative products; urban-rural price differences

1. INTRODUCTION

The International Comparisons Program (ICP) dates back to the 1960s. Its objective is to compare the purchasing power of currencies and real income across countries. ICP benchmarks, of which the most recent is for 2005, play a pivotal role in the construction of the Penn World Table.

ICP 2005 had a much larger budget than earlier rounds. This allowed far more data to be gathered for more countries (146 participated in ICP 2005).[1] Perhaps the most surprising result that emerged from ICP 2005 was that China came out about 40 percent smaller than previously thought (see Maddison, 2008; Chen and

Note: We thank the World Bank for funding this project and for providing the data, and Angus Deaton, Erwin Diewert, Yuri Dikhanov, Denzil Fiebig, Kevin Fox, Prasada Rao, two anonymous referees, and the Managing Editor Conchita D'Ambrosio for their comments. The views expressed here are those of the authors and do not necessarily represent those of the World Bank.

*Correspondence to: Robert J. Hill, Department of Economics, University of Graz, Universitätsstrasse 15/F4, 8010 Graz, Austria (robert.hill@uni-graz.at).

[1]Earlier rounds include 1970, 1973, 1975, 1980, 1985, and 1993. The number of participating countries increased gradually, from only 10 in 1970 to 117 in 1993.

Ravallion, 2010; Deaton and Heston, 2010; Feenstra *et al.*, 2013).[2] We return to this issue after some preliminary discussion of the mechanics of ICP 2005 and of the main innovations introduced in this paper.

The ICP 2005 aggregate results at the level of GDP are obtained from 155 basic heading price indexes and corresponding expenditure levels. A basic heading is the lowest level of aggregation at which expenditure data are available. A basic heading consists of a group of similar products defined within a general product classification. Food and non-alcoholic beverages account for 29 headings, alcoholic beverages, tobacco, and narcotics for 5 headings, clothing and footwear for 5 headings, etc. (see Blades, 2007a). The basic heading price indexes, which are typically calculated using the country–product–dummy (CPD) regression method (see Summers, 1973), together with their corresponding expenditure shares provide the building blocks from which the overall comparison is constructed. If these building blocks are biased or otherwise flawed, then everything that builds on them will be likewise tainted. Most errors are likely to arise in the process of calculating these basic heading price indexes and expenditure shares. It is here at this disaggregated level that the most pressing research problems can be found.[3]

In this paper we use a unique dataset consisting of over 600,000 price quotes drawn from nine countries in the Asia-Pacific region and covering 92 basic headings to explore ways of improving the quality of the basic heading price indexes in ICP. In ICP 2005 (and ICP 2011) the product prices for each country inserted into the CPD formula are arithmetic means of the individual price quotes. The individual price quotes are generally considered confidential and are not supplied to the ICP global office. The CPD model is estimated using ordinary least squares (OLS). We propose here a weighted least squares (WLS) version of CPD, which should potentially improve the efficiency of the estimated basic heading price indexes. To implement our WLS version of CPD, countries would need to supply geometric means instead of arithmetic means, the number of price quotes per product, and the standard deviation of the logs of the price quotes. Supplying these additional data should not violate any existing confidentiality restrictions. In the Appendix we also consider whether efficiency can be improved by estimating CPD-type models simultaneously over groups of basic headings.

We also try estimating CPD directly on the individual price quotes rather than on country average prices. While this might not be feasible in practice due to confidentiality restrictions, our dataset provides us with a unique opportunity to compare these two approaches. One attraction of estimating CPD on the individual price quotes is that it allows for the inclusion of urban–rural and outlet-type identifiers. The basic issue here is that a price index should compare like with like. For example, one should not compare directly the price of a product purchased in a rural area in an open market in country A with the price of the same product

[2]The World Bank's pre-ICP 2005 estimates for 2005 can be found in its World Development Indicators 2007 report (CD version). The printed version provides per capita gross national income (in table 1.1) rather than per capita GDP.

[3]Above basic heading level standard multilateral price index formulas such as GEKS or Geary–Khamis can be used. This higher level of aggregation has tended to attract much more attention in the literature (see, for example, Diewert, 1999; Hill, 2000; Neary, 2004).

bought in a supermarket in an urban area in country B. The inclusion of urban–rural and outlet-type identifiers provides a way of adjusting the results for these types of mismatches.

The basic CPD method can also be extended to include identifiers for representative products. Representative products are those that are widely available in a country and account for a significant proportion of total expenditure within a basic heading (see World Bank, 2008, p. 143). Unrepresentative products, meanwhile, are those products that are not widely available in an economy, and do not constitute a significant proportion of total expenditure within a basic heading. Other things equal, representative products tend to be cheaper than unrepresentative products. It follows that countries that sample disproportionately unrepresentative products—as may have been the case for China (see Deaton and Heston, 2010) in ICP 2005—will end up with price levels that are overestimated and per capita incomes that are underestimated. The inclusion of representative identifiers has the potential to correct such biases, although only if representative and unrepresentative products are identified in a reasonably consistent way across countries and the price difference between representative and unrepresentative products is sufficiently similar across products and countries.[4] The ICP 2005 Technical Advisory Group (TAG) indeed recommended that representative identifiers be included. In practice, only one region (South America) actually did so. Using our dataset we revisit this decision for the Asia-Pacific region.

Another contribution of our study is that we show how CPD-type methods can be used to measure urban–rural price differences across countries. This is an important topic in its own right that has implications for the poverty measurement literature. Our estimated urban–rural price differences also shed some light on the pre- and post-ICP 2005 discrepancy in the results for China. ICP 2005 has the advantage that it is a much more detailed comparison than the previous rounds. Pre-ICP 2005 results for China are obtained by extrapolation from a bilateral comparison between China and the U.S. in 1986 (see Rouen and Kai, 1995). It is tempting therefore to conclude that the problems lie with the pre-ICP 2005 results. However, China's participation in ICP 2005 was on a limited scale and the price quotes were obtained from only 11 cities and their surrounding areas (see Blades, 2007b). We conclude by exploring the extent to which the pre- and post-ICP 2005 discrepancy can be explained by China only sampling prices in urban areas.

The remainder of this paper consists of six sections. Section 2 explains the basic CPD model and some ways of extending it. The dataset is described in Section 3. Section 4 considers the question of whether the CPD model should be estimated using the individual price quotes or average prices, and if the latter, whether these average prices should be weighted or not? The estimated coefficients and the resulting basic heading and aggregate level price indexes derived from various versions of the CPD model are discussed in Section 5. Section 6 proposes some new CPD-based methods for measuring urban–rural price differences, and then applies these to our dataset. Some implications for China in ICP 2005 are also

[4]The same caution applies to the inclusion of urban-rural and outlet–type identifiers.

considered. Finally, Section 7 concludes the paper. Some further extensions of the CPD method are explored in an Appendix.

2. The Country–Product–Dummy Method and its Extensions

All regions in ICP 2005, with the exception of Eurostat–OECD, used a CPD-type method to calculate the within-region basic-heading price indexes for each country.[5] The CPD model estimates the following regression equation separately for each basic heading:[6,7]

$$\ln \bar{p}_{km} = \sum_{\mu=2}^{M} \alpha_\mu x_\mu + \sum_{j=1}^{K} \beta_j y_j + \varepsilon_{km}, \tag{1}$$

where p_{km} denotes the average of the price quotes for product m in country k, x_μ denotes a product dummy variable that equals 1 if $m = \mu$, and zero otherwise, while y_j denotes a country dummy variable that equals 1 if $k = j$ and zero otherwise, and ε_{km} denotes a random error term. Thus far in ICP the average price $\bar{p}_{km}$ used has been the arithmetic mean. Below we argue that it might be better to use geometric means.

The α_m and β_k parameters are typically estimated by ordinary least squares (OLS). Exponentiating the estimated β_k parameter, we obtain the price index p_k for this particular basic heading for country k, as follows:[8]

$$\hat{p}_k = \exp\left(\hat{\beta}_k\right).$$

In an ICP context, product m will only typically be available in a subset of the countries in the comparison. It is sufficient that m is priced in at least two countries for it to be included in the calculation of the basic heading price indexes.

The CPD model is estimated using OLS, which implies giving equal weight to the prices of all countries. If, however, we have reason to believe that the average prices from some countries are more reliable than from others, it may be more efficient to use WLS. Reliability can be measured by the variance of the log average prices.

We consider here two ways of calculating the variance when the average prices are arithmetic means.[9] First, suppose that the arithmetic mean of the price quotes

[5]One advantage of the CPD method is that its stochastic specification allows the use of a range of econometric tools and techniques that are not normally used in the computation of price indexes (see Rao, 2004). By contrast, Eurostat–OECD uses the nonstochastic EKS-S method to construct their basic heading price indexes (see Hill and Hill, 2009).

[6]It is common when estimating the CPD model to normalize the prices of one of the products and one of the countries to one. In this formulation, an additional constant term should be inserted in the equation. Here instead we omit the constant term but do not include a country normalization. Hence the summation over countries in (1) runs from $j = 1$ to K. The price of one product is still normalized to one, which is why the summation over products runs from $\mu = 2$ to M.

[7]In the Appendix we consider the possibility of estimating the CPD model simultaneously across multiple basic headings.

[8] $\exp\left(\hat{\beta}_k\right)$ is in fact a biased estimator of $exp(\beta_k)$. We find, however, that use of Kennedy's (1981) bias correction has virtually no impact on the results. Hence we ignore this correction here.

[9]We thank Prasada Rao for suggesting these methods.

for product m in country k, denoted by $\bar{p}_{km}^{A}$, is normally distributed with mean μ_{km} and variance $v(p_{km})/H_{km}$, where $v(p_{km})$ is the variance of the individual price quotes and H_{km} is the total number of price quotes obtained for product m in country k. It follows that $\ln(\bar{p}_{km}^{A})$ is log normal with the following variance:

$$v\left[\ln\left(\bar{p}_{km}^{A}\right)\right] = e^{[2\mu_{km}+v(p_{km})/H_{km}]}\left[e^{v(p_{km})/H_{km}} - 1\right].$$

Alternatively, a Taylor series approximation can be used as follows:

$$v[g(X)] \approx [g'(\mu_x)]^2 \, v(X).$$

In our case, this reduces to:

$$v\left[\ln\left(\bar{p}_{km}^{A}\right)\right] \approx \frac{v(p_{km})/H_{km}}{(\mu_{km})^2}.$$

Suppose now instead that geometric means are used. The variance of the log of the geometric mean price, denoted here by $v\left[\ln\left(\bar{p}_{km}^{G}\right)\right]$, can be calculated as follows:

$$(2) \qquad v\left[\ln\left(\bar{p}_{km}^{G}\right)\right] = v\left\{\ln\left[\prod_{h=1}^{H_{km}}\left(p_{km}^{h}\right)^{1/H_k}\right]\right\} = v\left[\frac{1}{H_{km}}\sum_{h=1}^{H_{km}}\ln\left(p_{km}^{h}\right)\right] = \frac{v[\ln(p_{km})]}{H_{km}},$$

where p_{km}^{h} denotes an individual price quote h on product m in country k.[10]

Irrespective of which approach is used, weights w_{km} for use in WLS can be calculated from the estimated variance of log average prices. For example, focusing on the geometric mean case, one way in which this can be done is as follows:

$$(3) \qquad w_{km} = \frac{1}{\bar{v}_m + v\left[\ln\left(\bar{p}_{km}^{G}\right)\right]},$$

where

$$\bar{v}_m = \frac{1}{K}\sum_{j=1}^{K} v\left[\ln\left(\bar{p}_{jm}^{G}\right)\right]$$

is the average variance of the log average prices across the whole set of countries. The term $\bar{v}_m$ is included to ensure that w_{km} is still defined when for a particular country k, $v\left[\ln\left(\bar{p}_{km}^{G}\right)\right] = 0$.

In summary, to calculate the weights for country–product km we need either arithmetic means $\bar{p}_{jm}^{A}$, variances of the price quotes $v(p_{km})$, and the number of price quotes H_{km}, or geometric means $\bar{p}_{jm}^{G}$, variances of the log price quotes $v[\ln(p_{km})]$ and the number of price quotes H_{km}. In our opinion geometric means are better since they rely on weaker assumptions. Either way though it should be possible for countries to provide these additional data without violating any confidentiality restrictions.

Returning again to OLS, an alternative extension of the CPD method—the country–product–representative–dummy (CPRD) method—was proposed by

[10]This derivation assumes only that the price quotes are identically and independently distributed—i.e., the covariances are zero and $\ln\left(p_{km}^{h}\right) = \ln\left(p_{km}\right)$.

Cuthbert and Cuthbert (1988). It simply adds an additional dummy variable z to the model as follows:

$$\ln \bar{p}_{km} = \sum_{\mu=2}^{M} \alpha_\mu x_\mu + \sum_{j=1}^{K} \beta_j y_j + \gamma z + \varepsilon_{km},$$

where z equals 1 if product m is representative in country k and zero otherwise.

The inclusion of representative dummies has the potential to correct the bias arising from price comparisons when a product is representative in one country and unrepresentative in another. However, the inclusion of representative dummies is justified only if representative products are uniformly cheaper than unrepresentative products in all countries. Furthermore, the extent to which representative products are cheaper should be the same across countries.

At its meeting in September 2004, the ICP 2005 Technical Advisory Group

> recommended that regions should use the CPRD method to estimate basic heading PPPs. Of course, the method can only be implemented satisfactorily if the countries within a region are able to identify representative products correctly. (Hill, 2007)

Hence all participating countries were asked to identify which of the products they priced were representative. However,

> Economies in the Asia-Pacific, Africa, Western Asia, and South America regions that either had not participated in an international comparison for an extended period or had never participated had difficulty applying the representativity concept, therefore, it was not used in their intraregional comparisons. (World Bank, 2008, p. 185)

It turns out this statement is not quite correct since South America did in fact use CPRD (see Diewert, 2008). It is true though that the Asia-Pacific region used CPD.

As has been noted above, in ICP 2005 (and ICP 2011), $\bar{p}_{km}$ is an average of the price quotes on product m obtained in country k. An alternative approach would be to include all the individual price quotes for product m directly in the CPD or CPRD regression. We would then have multiple observations on the price of product m in country k. The CPRD method can be further extended, when the individual price quotes are used, to include urban and outlet type dummies (i.e., the country–product–representative–urban–outlet–dummy (CPRUOD) method] as follows:

$$\ln \bar{p}_{km} = \sum_{\mu=2}^{M} \alpha_\mu x_\mu + \sum_{j=1}^{K} \beta_j y_j + \gamma z + \delta w + \sum_{i=2}^{I} \theta_i u_i + \varepsilon_{km}, \tag{4}$$

where now we also include a dummy w that equals 1 if product m is from an urban area in country k and zero otherwise, while $i = 1, \dots, I$ indexes a series of outlet types (supermarket, department store, open market, etc.). u_i is a dummy variable that equals 1 only if product m in country k was bought in an outlet of type i.

Using our dataset, we assess the feasibility of using the CPRUOD model and its country–product–urban–dummy (CPUD) and country–product–representative–urban–dummy (CPRUD) variants in an ICP context.

3. The Dataset

Our dataset consists of 605,998 price quotes for 2005 from the following nine countries in the Asia-Pacific region: Bhutan, Fiji, Hong Kong, Indonesia, Macao, Malaysia, the Philippines, Sri Lanka, and Vietnam.[11] In total there are 142 basic headings in ICP 2005 for the Asia-Pacific region (some other regions used 13 additional headings). Our price quotes are drawn from 92 of these headings, all of which belong in the Final Consumption Expenditure by Households category.[12] Our list of basic headings is shown in Table 1.[13]

For our purposes the dataset, while large, has some problems. Two countries (Hong Kong and Malaysia) identified all products as representative. Fiji did not identify any products as unrepresentative (although it did identify some as representative), while Vietnam failed to identify any products as either representative or unrepresentative. More generally, it seems likely that representativity was not identified in a consistent way across countries. The fact that three of the nine countries identified all products as representative is symptomatic of this lack of consistency. It is important that countries are provided with more guidance on this issue in future rounds of ICP.

Similarly, only six countries (Fiji, Indonesia, Malaysia, the Philippines, Sri Lanka, and Vietnam) supplied urban/rural identifiers. All the price quotes from Fiji are urban. Our biggest problems, however, related to the outlet-type data. As many as 41 different outlet types are identified in our data. However, it is impossible to match outlets across countries at this level of detail. We settled on sorting the outlet types into six groups: (i) department stores; (ii) supermarkets; (iii) open markets/stalls; (iv) specialized shops (traditional outlets); (v) wholesale and discount stores; and (vi) other stores. Some summary information is provided in Table 2. Further discussion on the quality of data is provided in Section 5.2.[14]

[11]Strictly speaking we should refer to economies rather than countries, given that two of our sample (Hong Kong and Macao) are not countries. Nevertheless, for convenience we will henceforth use the term "countries".

[12]In fact, we began with 95 basic headings. Our base country in all our comparisons is Hong Kong (Hong Kong is also the base in the official ICP 2005 comparisons for the Asia-Pacific region). Given that no data are available for Hong Kong for three headings, we decided therefore to exclude these from the comparison. This reduces the number of price quotes in our dataset from 610,024 to 605,998.

[13]Two of the most important and hard-to-measure headings in household consumption—namely consumption on rents of owner-occupiers and financial services indirectly measured (FISIM)—unfortunately are excluded. See, for example, Deaton (2005) for a discussion of these two headings and their potential impacts on international comparisons.

[14]A number of other outlet types were represented in the data (often sparsely and only for a small subset of countries). These included the following: minimarkets, kiosks and neighborhood shops; mobile shops and street vendors; other kinds of trade (mailorder, internet, etc); agencies; bakery; bank; Book store; bowling center; cinema; communication services; communication shop; computer shop; courier services; food court; furniture shop; gymnasium; holiday agencies; hotel; insurance agencies; motor vehicle outlet; music store; newspaper advertising; nursery; pet shop; petrol kiosk; photo kiosk; saloon; services outlet; shoe repair outlet; sundry shop; swimming pool; transportation services; pharmacy/drugstore; private doctor's clinic; public/government doctor's clinic; private hospital; public/government hospital; private dental clinic; public/government dental clinic; private laboratory; public/government laboratory; private optical clinic; puublic/government optical clinic; private outlet for therapeutic, appliances and equipment; public/government clinic for physiotherapist; private primary school; private secondary school; private college/university; private tutor.

TABLE 1

Our List of ICP Basic Headings for Final Consumption Expenditure by Households

1 110111.1	Rice	47 110531	Major household appliances
2 110111.2	Other cereals and flour	48 110532	Small electric household appliances
3 110111.3	Bread	49 110533	Repair of household appliances
4 110111.4	Other bakery products	50 110540	Glassware/tableware utensils
5 110111.5	Pasta products	51 110552	Small tools and misc. accessories
6 110112.1	Beef and veal	52 110561	Non-durable household goods
7 110112.2	Pork	53 110562.1	Domestic services
8 110112.3	Lamb, mutton and goat	54 110611	Pharmaceutical products
9 110112.4	Poultry	55 110612	Other medical products
10 110112.5	Other meats and meat prep	56 110613	Therapeutical appliances and equip
11 110113.1	Fresh, chilled or frozen fish	57 110621	Medical services
12 110113.2	Preserved or processed fish	58 110622	Dental services
13 110114.1	Fresh milk	59 110623	Paramedical services
14 110114.2	Preserved milk and milk products	60 110711	Motor cars
15 110114.3	Cheese	61 110712	Motor cycles
16 110114.4	Eggs and egg-based products	62 110713	Bicycles
17 110115.1	Butter and margarine	63 110722	Fuels/lubricants for transport equip
18 110115.3	Other edible oils and fats	64 110723	Maintenance of transport equipment
19 110116.1	Fresh or chilled fruit	65 110731	Passenger transport by railway
20 110116.2	Frozen, or processed fruit	66 110732	Passenger transport by road
21 110117.1	Fresh or chilled vegetables	67 110733	Passenger transport by air
22 110117.2	Fresh or chilled potatoes	68 110734	Passenger transport by sea/waterway
23 110117.3	Frozen or processed vegetables	69 110736	Other purchased transport services
24 110118.1	Sugar	70 110810	Postal services
25 110118.2	Jams, marmalades and honey	71 110820	Telephone and telefax equipment
26 110118.3	Confectionery, chocolate, ice	72 110830	Telephone and telefax services
27 110119	Food products n.e.c.	73 110911	Audio-visual/photographic equip
28 110121	Coffee, tea and cocoa	74 110914	Recording media
29 110122	Mineral waters, juices	75 110915	Repair of audio-visual/photo equip
30 110211	Spirits	76 110921	Durables for outdoor/indoor recreation
31 110212	Wine	77 110931	Other recreational items and equip
32 110213	Beer	78 110933	Gardens and pets
33 110220	Tobacco	79 110935	Veterinary and other services for pets
34 110311	Clothing materials	80 110941	Recreational and sporting services
35 110312	Garments	81 110942	Cultural services
36 110314	Cleaning, repair of clothing	82 110950	Newspapers, books and stationery
37 110321	Shoes and other footwear	83 110960	Package holidays
38 110322	Repair and hire of footwear	84 111000	Education
39 110430	Maintenance/repair of dwelling	85 111110	Catering services
40 110441	Water supply	86 111120	Accommodation services
41 110451	Electricity	87 111211	Hairdressing salons
42 110452	Gas	88 111212	Appliances/products for personal care
43 110453	Other fuels	89 111231	Jewellery, clocks and watches
44 110511	Furniture and furnishings	90 111232	Other personal effects
45 110512	Carpets and floor coverings	91 111262	Other financial services n.e.c
46 110520	Household textiles	92 111270	Other services n.e.c.

4. Average Prices Versus Individual Price Quotes in CPD-Type Regressions

Here we compare four variants on the CPD method: CPD(am), CPD(gm), CPD(wgm), and CPD(ipq). CPD(am) is the standard CPD method calculated using the arithmetic mean of the price quotes from each country. CPD(gm) differs from CPD(am) only in that it uses the geometric mean instead of the arithmetic mean. CPD(wgm) uses geometric means and is estimated using WLS with the weights calculated from (2). Finally, CPD(ipq) denotes CPD calculated using all the individual price quotes (ipq).

TABLE 2

Some Summary Information on Each Country

Countries	Outlet Type	Urban Price Guotes (%)	Rural Price Quotes (%)	Rep Price Guotes (%)	Unrep Price Quotes (%)	Number of Headings	Number of Price Quotes
Bhutan	Yes	100.0	0.0	59.8	16.4	74	17,085
Fiji	Yes*	100.0	0.0	18.8	0.0	70	9,897
Hong Kong	Yes	100.0	0.0	100.0	0.0	92	45,231
Indonesia	No	38.2	61.8	98.5	1.5	40	62,972
Macao	Yes	100.0	0.0	95.9	4.1	91	28,554
Malaysia	Yes	83.9	16.1	100.0	0.0	85	70,683
Philippine	Yes	83.1	16.9	92.2	7.8	85	142,379
Sri Lanka	No	58.2	41.8	53.3	7.3	84	72,562
Vietnam	No	57.9	31.7	100**	0**	83	156,635
Total		71.9	25.5	89.8	3.5		605,998

Notes: Urban/rural identifiers are missing for 10.4 percent of price quotes in Vietnam. Rep/untrep identifiers are missing for 23.8, 81.2, and 39.4 percent of price quotes in Bhutan, Fiji, and Sri Lanka respectively.

*Outlet type identifiers are missing for many of Fiji's price quotes.

**Vietnam did not provide any rep/unrep identifiers. We have assumed that all Vietnam's price quotes are representative.

CPD(ipq) has two advantages over average price CPD methods. First, using the individual price quotes dramatically increases the degrees of freedom in the CPD model. Second, it allows the inclusion of outlet or urban–rural dummies. It also has three disadvantages. First, the individual price quotes are considered sensitive information in some countries and regions (e.g., the European Union). Indeed in some cases countries are unable to provide individual prices quotes because of their national laws governing data confidentiality. Hence the ICP global office will typically not be given access to these data. Second, a country that has more price quotes for a particular product may exert more influence on the overall quality adjustment factor for that product (for a detailed proof, see Diewert, 2004). This is potentially problematic in an ICP context in which all countries are supposed to be treated symmetrically. The average price method, by contrast, is democratic in the sense that each country has roughly the same influence on the determination of the quality adjustment factor for a basic heading category (again, see Diewert, 2004). Third, the inclusion of urban or outlet-type dummies can themselves introduce biases into the comparison (see Section 5.4).

It should be possible to obtain the weights w_k required to implement WLS without violating confidentiality rules—for the case of geometric means, see (2) and (3). Also, the use of weights has the potential to increase the efficiency of the estimated basic heading price indexes. One potential problem with CPD(wgm) is that sometimes a low or even zero value for the variance of the log price quotes in a country k (i.e., $v[\ln(p_k)]$ in (2)), far from signaling a high level of precision in price measurement may in fact imply quite the opposite. In particular, if there is only a single price quote from country k then $v[\ln(p_k)]$ necessarily equals zero. Also, for hard to measure headings, such as 40-Water supply, where no genuine price quotes are available, a country may simply set a reference price which is the same in all locations. Hence again $v[\ln(p_k)]$ equals zero. One solution in such cases is to assume that $v[\ln(p_k)]$ is the same for all countries, and then focus exclusively on the number of price quotes as the signal of reliability. Alternatively, the use of WLS can be restricted to basic headings where these kinds of problems are deemed unlikely to arise.

The four variants on CPD (i.e., CPD(am), CPD(gm), CPD(wgm), and CPD(ipq)) are compared in Table 3 using the following metric:

$$A_k(x, y) = \frac{100}{N(K-1)} \sum_{b \neq k}^{K} \sum_{n=1}^{N} \left[\max\left(P^x_{bn,kn} \big/ P^y_{bn,kn},\, P^y_{bn,kn} \big/ P^x_{bn,kn} \right) - 1 \right]. \tag{5}$$

The metric compares a pair of methods, denoted by x and y. For example, x could be CPD(am), while y could be CPD(gm). N is the number of basic headings and K the number of countries in the comparison. $P^x_{bn,kn}$ denotes the price index of country k for basic heading n calculated using method x, expressed relative to the corresponding price index of the base country b. The comparison is made by using each of the other countries in turn as the base.

The metric $A_k(x, y)$ measures the average percentage difference between the basic heading price indexes generated by methods x and y for country k. For example, in Table 3, $A_k(x, y) = 4.3$ for Indonesia when the methods being compared are CPD(am) and CPD(gm). This means that by switching from CPD(am) to CPD(gm), the basic heading price indexes for Indonesia change by on average 4.3 percent.

From Table 3 it can be seen that the impact of switching from using arithmetic means to geometric means is relatively small. The average difference between the CPD(am) and CPD(gm) basic headings is 2.9 percent. The switch to WLS has a bigger impact. The CPD(am) and CPD(wgm) headings differ on average by 8.7 percent. This difference is quite large, implying that the choice between CPD(am) and CPD(wgm) is a matter of empirical significance. The average difference between CPD(wgm) and CPD(ipq) is 6.2 percent, which is less than the average difference of 8.7 percent between CPD(am) and CPD(wgm). In this sense CPD(wgm) is closer to CPD(ipq) than to CPD(am). This suggests that WLS may provide some of the advantages of CPD(ipq) without the associated disadvantages (e.g., violations of confidentiality restrictions).

We also compare the four variants on CPD (i.e., CPD(am), CPD(gm), CPD(wgm), and CPD(ipq)) using a second metric. This second metric computes the price level dispersion across countries at basic heading level. The price levels are obtained here by dividing each country k's price index P_{kn} for basic heading n by its corresponding average 2005 market exchange rate MER_k, with Hong Kong in both cases normalized to 1. Dispersion is then measured by taking the standard deviation of the log price levels as follows:[15]

$$(6) \qquad \sigma_n = \sqrt{\sum_{k=1}^{K} \frac{\left[\ln(P_{kn}/MER_k) - \overline{\ln(P_{kn}/MER_k)}\right]}{K-1}},$$

where $\overline{\ln(P_{kn}/MER_k)}$ is the average log price level for basic heading n. The relative magnitudes of price level dispersion across basic headings for a pair of methods (again denoted by x and y) are compared in Table 4. For example, in a comparison between CPD(am) and CPD(gm), σ_n is larger for CPD(am) for 33 basic headings, while for the remaining 59 headings, σ_n is larger for CPD(gm). The final row of Table 4 provides standard normal test statistics Z derived from the normal approximation to the binomial distribution (with mean $N/2$ and variance $N/4$, where N is the number of basic headings). The Z statistics indicate that the differences in price level dispersion at basic heading level are significant at the 95 percent level for CPD(am) versus CPD(gm) and for CPD(wgm) versus CPD(ipq). These findings are somewhat at odds with the results in Table 3. This may be because the approach in Table 4 fails to take account of the magnitudes of the differences between σ_x and σ_y, or the relative importance of the basic headings (i.e., some have larger expenditure shares than others). A better indication of how important are differences in price level dispersion across CPD-type methods is obtained by comparisons at the aggregate level (here Household Consumption). We return to this issue in Section 5.3.

In summary we believe that the WLS approach described above deserves serious consideration in future rounds of ICP. The difference between CPD(am) and CPD(wgm) is not trivial, and the CPD(wgm) estimated basic heading price indexes should be more efficient. The use of CPD(wgm) would require, however, a slight change in the summary data for the price quotes provided by participating

[15]Taking logs before computing the standard deviation ensures that the results are invariant to the choice of base country.

TABLE 3

Average Differences A_K in the Basic Heading Price Indexes Between Pairs of Methods

	CPD(am) CPD(gm)	CPD(am) CPD(wgm)	CPD(am) CPD(ipq)	CPD(gm) CPD(wgm)	CPD(gm) CPD(ipq)	CPD(wgm) CPD(ipq)	CPD CPRD	CPD CPRUD	CPRD CPRUD
BHU	3.1666	10.1883	17.5496	8.0430	15.2048	7.4204	10.7592	10.9272	1.1371
FIJ	5.3848	11.6015	17.0862	7.1345	12.4943	6.7028	24.7157	24.8598	1.0361
HKG	2.3641	7.9759	13.2306	7.0599	12.2122	5.9790	6.3058	6.6406	0.9791
INO	4.2874	10.9936	18.0294	9.8299	16.7009	7.1467	4.7454	5.6347	1.9624
MAC	2.3228	7.4332	12.4706	6.3227	11.0515	6.1546	6.4628	6.7075	0.9798
MAL	2.1446	8.0952	14.2264	7.3564	13.4491	6.3546	6.2314	6.4916	0.9524
PHI	2.2755	7.8192	13.2095	7.3014	12.6168	5.6839	7.0184	7.2337	0.9330
SRI	2.1526	7.1176	11.3026	6.1701	10.2091	5.4738	10.2593	10.4051	1.4188
VIE	2.1331	6.7123	11.1632	5.7895	10.0008	5.1620	6.2197	6.7468	1.5111
Aver	2.9146	8.6597	14.2520	7.2231	12.6600	6.2309	9.1909	9.5163	1.2122

Notes: The CPD(am), CPD(gm), and CPD(wgm) basic heading price indexes are calculated using average prices, while CPD(ipq), CPRD, and CRPUD are calculated using the individual price quotes.

TABLE 4

A Comparison of Price Level Dispersion at Basic Heading Level Across Methods

x y	CPD(am) CPD(gm)	CPD(am) CPD(wgm)	CPD(am) CPD(ipq)	CPD(gm) CPD(wgm)	CPD(gm) CPD(ipq)	CPD(wgm) CPD(ipq)	CPD CPRD	CPD CPRUD	CPRD CPRUD
σx > σy	33	46	42	47	40	34	42	43	55
σx < σy	59	46	50	41	48	54	48	47	29
Z	−2.71	0.00	−0.83	0.64	−0.85	−2.13	−0.63	−0.42	2.84

Notes: σx denotes the standard deviation of the country price levels for a particular basic heading (calculated using method x). For a pair of methods (say CPD and CPRD) we count how many basic headings have smaller standard deviations for the CPD method (denoted by ax) than for the CPRD method (denoted by σy). The total number of basic headings available depends on the pair of methods being compared. The Z values are derived from the normal approximation to the binomial distribution based on the null hypothesis that the probability that σx > σy is 0.5.

countries. Instead of providing arithmetic means, they would need to provide geometric means, the number of price quotes, and the standard deviation of log prices for each basic heading.

5. Including Representative, Urban, and Outlet-Type Dummy Variables in a CPD-Type Model

5.1. *The Plausibility of the Estimated Coefficients on the Representative, Urban, and Outlet-Type Dummy Variables*

Confidentiality issues notwithstanding, here we try running an extending CPD model on the individual price quotes that includes urban and outlet-type dummies in addition to representative dummies.[16] As far as we are aware, this has never been tried before. Our objective is to assess the plausibility of the estimated representative, urban, and outlet-type coefficients.

We assume in what follows that all prices in Vietnam are representative and that all prices in Bhutan, Hong Kong, and Macao are urban. Even so, not all countries can be included in all 92 basic heading regressions. For example, Indonesia provided data only for 41 headings. Hence it is excluded from 51 of our basic heading regressions.

Some summary statistics from our estimated CPRUOD models are shown in Table 5. Here we focus on the signs of the estimated representative, urban, and outlet-type coefficients. Taking the representative coefficients first, our prior expectation is that the sign of these coefficients should be negative. That is, other things equal, representative products should be cheaper than unrepresentative products. The results are only weakly supportive of this hypothesis; 42 coefficients are negative and 35 are positive. Of the statistically significant coefficients at the 5 percent level, 27 are negative and 21 positive. Our prior for the urban coefficients is that they should be positive since, other things equal, prices tend to be higher in urban areas than in rural areas. The situation in this case, however, is not so clear cut since there may be exceptions to this rule. For example, imported products may be more expensive in rural areas due to greater transport costs and less competition amongst retailers. The results broadly support our hypothesis, with 54 coefficients being positive (and 33 statistically significant) and only 26 being negative (with 11 statistically significant).[17]

The priors for outlet type are less obvious. Other things equal, it seems plausible that prices should be higher in "department stores" than in "supermarkets," and prices in "supermarkets" should be higher than in "open markets" and "wholesale discount stores." Given the heterogeneity of the "specialized stores" and "other store" categories, it is difficult to form any priors on them. The base outlet type is "supermarkets." The "department stores" coefficient is positive for 28 headings (11 of which are significant) and negative for 27 coefficients (10 of which are significant). Hence there is no discernible pattern here. The results are more

[16]Representative dummies can be included in a CPD-type model irrespective of whether CPD is run using average or individual prices.

[17]The total number of headings covered changes depending on whether our focus is on representative, urban, or outlet-type dummies since these identifiers are not available for all headings.

TABLE 5

Some Statistics on the Signs and Significance Levels of the Estimated Coefficients of the CPRUOD Model

Variables	Statistics	All	Positive	Negative
Representative variable				
	Number of +ve/–ve sign coefficients		35	42
	Number of significant coefficients		21	27
	Simple average of coefficients	–0.1	0.148	–0.3
Urban variable				
	Number of +ve/–ve sign coefficients		54	26
	Number of significant coefficients		33	12
	Simple average of coefficients	0.018	0.075	–0.1
Outlet-type variables*				
Department stores	Number of +ve/–ve sign coefficients		28	27
	Number of significant coefficients		11	10
	Simple average of coefficients	–0.026	0.144	–0.201
Open markets	Number of +ve/–ve sign coefficients		32	47
	Number of significant coefficients		12	23
	Simple average of coefficients	–0.031	0.133	–0.143
Specialized stores	Number of +ve/–ve sign coefficients		27	60
	Number of significant coefficients		14	43
	Simple average of coefficients	–0.047	0.165	–0.143
Wholesale & discount stores	Number of +ve/–ve sign coefficients		12	22
	Number of significant coefficients		6	12
	Simple average of coefficients	–0.069	0.169	–0.198
Other stores	Number of +ve/–ve sign coefficients		36	56
	Number of significant coefficients		12	38
	Simple average of coefficients	0.005	0.139	–0.097

Note: *The base outlet type is supermarkets.

plausible for "open markets" and "wholesale and discount stores" (i.e., they are both cheaper than "supermarkets") although still very noisy. For open markets, 47 coefficients are negative (of which 23 are significant), while 32 are positive (of which 12 are significant). For discount stores, 22 coefficients are negative (of which 12 are significant), while 12 are positive (of which 6 are significant).

We suspect that there may be serious inconsistencies with the way that outlet types are identified across countries, and that this may explain the erratic results. We would recommend that in future rounds of ICP the range of outlet types be significantly reduced. The six we consider might constitute a useful starting point. Also, it is important that these six categories are interpreted in a consistent way across countries. For example, it seems from the current results that the term "department store" may not mean the same thing in all nine countries in our dataset.

For these reasons, we henceforth exclude outlet-type dummies from our regression model. Our focus now is the CPRUD model. The results are presented in Table 6. The signs of the representative coefficients here accord a bit better with our prior expectations, with 48 negative coefficients (of which 43 are significant) and 29 positive coefficients (of which 20 are significant). This is in spite of the fact that Hong Kong and Malaysia identified every single product as representative (a clear sign that this terminology was not interpreted in a consistent way across countries). The coefficient on the urban dummy is typically positive as expected,

TABLE 6

Some Statistics on the Signs and Significance Levels of the Estimated Coefficients of the CPRD and CPRUD Models

Model	Variable/S	All	Positive	Negative
CPRD model	*Representative variable*			
	Number of +ve/–ve sign coefficients		30	47
	Number of significant coefficients		20	35
	Simple average of coefficients	–0.123	0.145	–0.294
CPRUD model	*Representative variable*			
	Number of +ve/–ve sign coefficients		29	48
	Number of significant coefficients		20	43
	Simple average of coefficients	–0.123	0.148	–0.287
	Urban variable			
	Number of +ve/–ve sign coefficients		63	21
	Number of significant coefficients		45	14
	Simple average of coefficients	0.026	0.052	–0.053

being 63 times positive (of which 45 are significant) and 21 times negative (of which 14 are significant). Also, shown in Table 6 are results for the CPRD method. The results for CPRD are similar to those obtained for the representative dummies in CPRUD.

The inclusion of representative dummies presupposes a consistent difference in the price of representative and unrepresentative products within a basic heading across all countries. Given that Hong Kong and Malaysia identified all products as representative, while Vietnam left this column blank, it is therefore far from clear that the inclusion of representative dummies would have improved the results for the Asia-Pacific region in ICP 2005. In particular, the use of CPRD in this context would actually cause an upward bias in the resulting price indexes for Hong Kong and Malaysia (assuming that the classification of all products as representative in these countries was erroneous). Hence we are inclined to agree with the decision to use CPD in preference to either CPRD or CPRUD for the Asia-Pacific region in ICP 2005.[18] A stronger case can perhaps be made for the inclusion of urban dummies (i.e., the CPUD model). However, this also can create problems (see Section 5.4).

It is evident from the above that while the dataset is large and unique, it suffers from some serious drawbacks. It demonstrates some of the challenges that ICP faces in arriving at its results. Our analysis of the data shows that much can be improved in price comparisons with regard to data collection, harmonization of important definitions and concepts across countries, and sampling techniques. Attaining this objective will require continued effort from ICP and, perhaps more importantly, the participating countries.

5.2. *Differences in Estimated Basic Heading Price Indexes Across Methods*

While the discussion above indicates that for our dataset CPD should probably be preferred to either CPRD or CPRUD, it is still interesting to consider how

[18]In recognition of these problems, ICP is asking participating countries to identify important (in terms of expenditure share) rather than representative products in the next round (i.e., ICP 2011).

much the choice of method affects the resulting basic heading price indexes. The average differences in the basic heading price indexes, as measured by the A_k metric defined in (5), of CPD versus CPRD, CPD versus CPRUD, and CPRD versus CPRUD (all computed over the individual price quotes) are shown in Table 3.[19]

From Table 3 it can be seen that the differences between CPD and CPRD are quite large (on average 9.2 percent). A few factors are probably contributing to this finding. First, the coverage of basic headings differs significantly across countries (as shown in Table 2). Indonesia, for example, only provides data on 41 headings. Hence the low value of its A_k(CPD,CPRD) coefficient can be attributed largely to its complete omission of many of the more problematic headings. Second, often larger values of $A_k(x, y)$ may be attributable primarily to differences in the underlying datasets rather than the methods themselves. For example, representative–unrepresentative indicators are available for only 22 percent of price quotes in Fiji. It follows that the CPRD results for Fiji are calculated on a much smaller dataset than the corresponding CPD results. Third, for ten headings the CPRD and CPRUD models were not identified. For seven of these cases data were only available for Hong Kong and Macao, and all the price quotes were representative and urban. For these headings, we set the CPRD and CPRUD results equal to the CPD results. For two other headings (40–Water supply and 41–Electricity) all the price quotes were representative, although there were both urban and rural price quotes. In these cases it was possible to estimate the CPUD but not the CPRD or CPRUD model. For these headings we set CPRD equal to CPD and CPRUD equal to CPUD. Finally, for basic heading 75 (Repair of audio-visual, photographic, and information processing equipment) all price quotes were representative for all countries except Macao, where all price quotes were unrepresentative. In this case again CPRD is set equal to CPD, and CPRUD is set equal to CPUD. These substitutions may cause the A_k coefficients to underestimate the underlying sensitivity of the results to the choice of method (although this effect is likely to be swamped by the effect of unmatched samples across methods discussed above).

In a comparison between CPD and CPRD, the biggest changes are observed for Fiji, where the results on average change by 25 percent. As noted above, most of this change is probably attributable to the large differences in the datasets used to calculate the CPD and CPRD results, rather than inherent differences in the underlying methods.

For headings where a switch from CPD to CPRD causes a large fall in the number of usable price quotes, any gains from the additional information provided by the inclusion of representative dummies will probably be outweighed by the loss of information caused by the exclusion of price quotes for which representative–unrepresentative indicators are not available. An important implication of this insight is that even if CPRD was adopted in the next round of ICP, it would still be preferable to use CPD for headings where the representative–unrepresentative indicators are particularly sparse. The same principle applies for CPRUD and CPRUOD. These methods should not be applied uniformly to

[19]We tried also estimating the CPD, CPRD, and CPRUD models using feasible generalized least squares (FGLS) to correct for heteroskedasticity. This had little impact so we do not present these results here.

all headings. More generally, we can imagine a future scenario where CPRUOD is used for one group of headings, CPRUD for a second group, CPRD for a third group, and finally CPD for a fourth group of particularly problematic headings.

We now turn to the issue of whether there are systematic differences in basic heading price level dispersion for CPD, CPRD, and CPRUD methods. For example, we find that σ_n, as defined in (6), is higher for the CPRUD method than for CPRD for 29 headings and lower for 55 headings, as shown in Table 4.[20] Hong Kong and Malaysia identified all products as representative (and we assumed that Vietnam's price quotes were all representative). This should cause the price levels in these countries to be higher relative to the other countries under CPRD than under CPD. It is not clear though how this impacts on overall price level dispersion since Hong Kong is a high price level country, Malaysia a middle price level country, and Vietnam a low price level country (see Table 7). It is therefore not surprising that the switch from CPD to CPRD does not have a statistically significant impact on overall price level dispersion in Table 4 as measured by the test statistic Z (obtained from the normal approximation to the binomial distribution).

By contrast, in a comparison of CPRD with CPRUD, the CPRD σ_n coefficient is higher for 55 headings, and smaller for only 29 headings. In this case $Z = 2.837$, which is significant at the 5 percent level. This is probably because the three countries with highest overall price levels (see Table 7)—namely Fiji, Hong Kong, and Macao—only supply urban prices. The inclusion of urban dummies therefore acts to lower slightly the relative price levels in these three countries, thus reducing overall price level dispersion.

5.3. *Results at the Level of Household Consumption*

Each of our CPD-type methods generates its own set of basic heading price indexes. Combining these with the corresponding expenditure shares obtained form ICP 2005, it is possible to compute Fisher price indexes at the level of Household Consumption for each CPD-type method. Here we use Hong Kong as the base. We use Fisher in preference to Gini–Eltetö–Köves–Szulc (GEKS) (see Gini, 1931; Eltetö and Köves, 1964; Szulc, 1964) since the list of basic headings available differs from one country to the next. Hong Kong however has data on all 92 headings. Hence comparing each country with Hong Kong maximizes the list of headings over which the Fisher price indexes are calculated.[21]

The resulting price levels obtained by dividing the Fisher price indexes by market exchange rates are shown in Table 7 (with Hong Kong normalized to 1). The methods considered in Table 7 are CPD(am), CPD(gm), CPD(wgm), CPD(ipq), CPRD, and CPRUD. The final row of Table 7 shows the price level dispersion of each method (measured by the standard deviation of the log price levels). It is noticeable that the CPRD and CPRUD methods have significantly

[20]As was noted above, for seven headings, only Hong Kong and Macao supplied data and for these headings all products were representative and urban. Hence it follows that there is no difference between the CPD and CPRD models in these cases. Hence we are left with 85 usable headings in a CPD–CPRD comparison. In addition there is another heading for which all products priced by participating countries were representative and urban, further reducing the usable headings to 84.

[21]It follows that the results for each bilateral comparison do not cover all of household consumption.

TABLE 7

Price Levels for Household Consumption (HKG = 1)

	CPD(am)	CPD(gm)	CPD(wgm)	CPD(ipq)	CPRD	CPRUD
BHU	0.5107	0.5055	0.4978	0.4871	0.4700	0.4700
FIJ	0.8487	0.8398	0.7994	0.8093	0.9318	0.9332
HKG	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
INO	0.2950	0.2785	0.2816	0.2757	0.2727	0.2767
MAC	0.8739	0.8808	0.8712	0.8646	0.8680	0.8674
MAL	0.6547	0.6595	0.6510	0.6498	0.6528	0.6545
PHI	0.5266	0.5243	0.5074	0.5048	0.5057	0.5072
SRI	0.4934	0.4936	0.4872	0.4751	0.4705	0.4763
VIE	0.4650	0.4612	0.4714	0.4747	0.4785	0.4845
σx	0.3838	0.3984	0.3897	0.3973	0.4200	0.4153

Notes: Price levels are calculated by dividing the Fisher price index by the market exchange rate (with HKG normalized to 1). For example. according to CPD(am), the price level in Bhutan is only 51 percent of that in Hong Kong, σx is the standard deviation of the logs of the price levels. It measures the level of price level dispersion.

higher price level dispersions σ_x than the other methods. The main driver of this result is Fiji, whose price level rises from 0.81 to 0.93 relative to Hong Kong as a result of switching from CPD to CPRD or CPRUD. Fiji is problematic for two reasons. First, as shown in Table 2, it has by far the least price quotes (i.e., only 9897). Second, over 80 percent of these price quotes are deleted when we switch from CPD to CPRD due to the lack of representativity identifiers. This second factor is presumably driving the big changes in the results from CPD to CPRD.

A comparison of CPD(am), CPD(gm), CPD(wgm), and CPD(ipq) reveals that the results at the aggregate level are not as sensitive to the choices between these methods as are the results at basic heading level (shown in Table 3). This is presumably because some of the differences observed at basic heading level offset each other at the aggregate level. Nevertheless, a comparison between the current ICP method—CPD(am)—and our preferred method—CPD(wgm)—still reveals some important differences. For example, a change in the Philippines price level from 0.53 to 0.51 (relative to Hong Kong) is not trivial.

5.4. *Correcting for Differences in the Urban–Rural Expenditure Mixes Across Countries in CPUD-Type Models*

Hong Kong is 100 percent urban in terms of both its price quotes and population. CPUD-type methods tend to exert downward pressure on the observed price level for Hong Kong as a result of all its price quotes being identified as urban. Such an adjustment may not be justified since households in Hong Kong do not have the option of purchasing in rural areas (without traveling beyond its borders). At the heart of this is the following question.

> Suppose countries *j* and *k* sell exactly the same products at exactly the same prices (converted at market exchange rates). However, country *j* is predominantly urban while country *k* is predominantly rural. Should these two countries have the same price level?

We assume that most users would say the answer is "yes." According to the CPUD method, however, the answer is that the price level is higher in the predominantly rural country k.

The problem with CPUD is that it implicitly assumes that the expenditure mix across urban and rural areas is the same in all countries. Hence to prevent bias when these expenditure mixes differ, an adjustment is required. Let Exp_{Urb}^{k} and Exp^{k} denote urban and total expenditure, respectively, in country k. One possible way of adjusting CPUD basic heading price indexes is as follows:

$$\tilde{P}_n^k = \left[\left(\frac{\text{Exp}_{Urb}^{k}}{\text{Exp}^{k}}\right)(P_{Rur,Urb} - 1) + 1\right] P_n^k, \tag{7}$$

where P_n^k denotes the original CPUD price index for basic heading n in country k, $\tilde{P}_n^k$ is the adjusted index, and $P_{Rur,Urb}$ is the average CPUD urban–rural price differential derived from (10) below.[22] From (7) we can see for a predominantly urban population that $\tilde{P}_n^k \approx P_{Rur,Urb} \cdot P_n^k > P_n^k$, while for a predominantly rural population $\tilde{P}_n^k \approx P_n^k$. That is, the price index of a predominantly urban country gets scaled up by almost the full urban–rural price differential while the price index of a predominantly rural country is left almost unchanged. More generally, the upward adjustment of a country's price index in (7) depends both on its urban share in total expenditure, and on the magnitude of the urban–rural price differential. When all countries have the same urban–rural expenditure mix, then all the price indexes get scaled up by the same factor, which effectively means they do not change (since price indexes are invariant to rescaling). That is, in this case the CPUD method gives the right answer.[23,24]

6. Measuring Price Differences Between Urban and Rural Areas

6.1. *Quality Unadjusted Urban-Rural Price Differences*

Poverty counts in large countries such as India and China can be highly sensitive to measured price differences between urban and rural areas. CPD-type methods can be used to shed light on this issue.

Here we consider three approaches to calculating urban–rural price differentials. The first and simplest is to take the ratio of the geometric means of the rural and urban price quotes in a particular heading for a particular country k.

$$\frac{P_{kn}^{Urb}}{P_{kn}^{Rur}} = \frac{\left(\prod_{u=1}^{U_n} p_{knu}^{Urb}\right)^{1/U_n}}{\left(\prod_{r=1}^{R_n} p_{knr}^{Rur}\right)^{1/R_n}}, \tag{8}$$

[22] With this adjustment, it will in general no longer be the case that the price index of one country is normalized to one. If such a normalization is desired, this can be achieved by dividing through the price indexes of all countries by the price index of the base country.

[23] Another option is to use weighted CPD where the weights reflect the urban–rural composition of the transactions for a product in a country.

[24] Similar problems can arise for CPRD-type methods.

TABLE 8

Urban–Rural Price Differentials

	Unweighted Geometric Mean	Weighted Geometric Mean	S.D. of Log Price Diff.
GM–Indonesia	1.022	1.093	0.377
GM–Malaysia	1.148	1.199	0.431
GM–Philippines	1.185	1.128	0.420
GM–Sri Lanka	1.044	1.055	0.112
GM–Vietnam	1.159	1.117	0.292
GM–Average	1.110	1.120	–
CPD–Indonesia	0.990	1.000	0.191
CPD–Malaysia	1.078	1.101	0.151
CPD–Philippines	1.008	1.003	0.137
CPD–Sri Lanka	1.024	1.034	0.061
CPD–Vietnam	1.030	1.038	0.077
CPD–Average	1.026	1.026	–
CPUD	1.027	1.029	0.053

Notes: The rural region and representative products are the numeraires, respectively. For example, a rural–urban price differential of 1.022 implies that urban prices are 2.2 percent higher than rural prices. GM–XXX denotes a price differential for country XXX calculated using either the unweighted or weighted formula in equation (9). CPD–XXX denotes a price differential for country XXX calculated using equation (11). CPUD is a price differential calculated using equation (10). Also shown in the final column are the standard deviations of the logarithms of the price differentials across all basic headings for each country.

where p_{knr}^{Rur} denotes rural price quote r and p_{knu}^{Urb} denotes urban price quote u for basic heading n in country k. Unweighted and weighted geometric averages of the urban–rural price differentials for each country can now be obtained as follows:

$$(9) \qquad \text{Unweighted: } \frac{P_k^{Urb}}{P_k^{Rur}} = \prod_{n=1}^{N}\left[\frac{P_{kn}^{Urb}}{P_{kn}^{Rur}}\right]^{1/N}, \quad \text{Weighted: } \frac{P_k^{Urb}}{P_k^{Rur}} = \prod_{n=1}^{N}\left[s_{kn}\frac{P_{kn}^{Urb}}{P_{kn}^{Rur}}\right],$$

where s_{kn} is the expenditure share of basic heading n in country k. The resulting unweighted and weighted average urban–rural price differentials for all countries for which we have rural and urban identifiers (i.e., Indonesia, Malaysia, Philippines, Sri Lanka, and Vietnam) are shown in Table 8. The overall average unweighted and weighted differentials are 11 and 12 percent, respectively (i.e., urban prices are 11 or 12 percent higher than rural prices).

6.2. *CPD-Based Approaches to Quality Adjusting Urban–Rural Price Differences*

One problem with the method described above is that it does not compare like with like. That is, the rural and urban price quotes are not matched to the same products. The CPUD method can be used to correct this problem. The CPUD regression model takes the following form:

$$(10) \qquad \ln \bar{p}_{km} = \kappa + \sum_{\mu=2}^{M}\alpha_\mu x_\mu + \sum_{j=2}^{K}\beta_j y_j + \delta w + \varepsilon_{km},$$

where m indexes the products in the basic heading, α and β are respectively the coefficients on the product and country dummies, and δ is the coefficient on the urban dummies. Estimating the CPUD model for each basic heading, we obtain 92 $\hat{\delta}$ coefficients. The exponent of each of these coefficients $\exp(\hat{\delta})$ can be interpreted as a price index measuring the average price difference between urban and rural areas, with rural as the numeraire, for a heading. Unweighted and weighted geometric means across all basic headings of the exponents of the δ coefficients are also provided in Table 8. The unweighted differentials give equal weight to each basic heading (and to each country in the overall average), while the weighted differentials weight each heading by its expenditure share in that country (and each country in the overall average by its share of total GDP calculated at official ICP purchasing power parity exchange rates).

One weakness of the CPUD method is that it assumes that the urban–rural price differential is the same for all countries. This is unlikely to be the case. For example, to the extent that price differentials are caused by transport costs, domestically produced food should be cheaper in rural areas where it is produced, while imported food should be cheaper in urban areas (e.g., ports). Hence countries that import more of their food may tend to have lower urban–rural price differentials than countries that produce most of their own food. In addition, concerns were raised in ICP 2005 that participating countries did not necessarily distinguish between rural and urban zones in a consistent manner (see Vogel, 2010).

This problem can be addressed using a variant on the standard CPD method that treats the rural and urban areas in each country as two separate entities as follows:

$$\ln \bar{p}_{km} = \kappa + \sum_{\mu=2}^{M} \alpha_{\mu} x_{\mu} + \sum_{j=2}^{K} \beta_j^{Rur} y_j^{Rur} + \sum_{j=2}^{K} \beta_j^{Urb} y_j^{Urb} + \varepsilon_{km}, \tag{11}$$

where y_j^{Rur} is a dummy that equals 1 only if that particular price quote is from a rural area in country *j*, while y_j^{Urb} equals 1 if the price quote is from an urban area in country *j*. The ratio $\exp(\hat{\beta}_j^{Urb}) / \exp(\hat{\beta}_j^{Rur})$ can be interpreted as a urban–rural price index for country *j* for that particular basic heading.[25] As shown in Table 8, both the unweighted and weighted average differentials now are 2.6 percent.[26] This finding suggests that the urban price quotes are drawn more from the expensive products within a basic heading while the rural price quotes are drawn more from the cheaper (presumably lower quality) products. If so, it follows that a simple ratio of average price quotes, due to its failure to quality adjust, overstates the actual price differential between rural and urban areas. When we quality adjust, we find that urban prices are only about 2.6 percent higher than rural prices.[27]

[25]We thank Angus Deaton for suggesting this method to us.

[26]Again, the unweighted differentials give equal weight to each basic heading (and to each country in the overall average), while the weighted differentials weight each heading by its expenditure share in that country (and each country in the overall average by its share of total GDP calculated at official ICP purchasing power parity exchange rates).

[27]Our finding that quality adjustment reduces the urban–rural price differential by about 8 percentage points is consistent with a similar finding by Deaton and Dupriez (2011). Using unit value data obtained from India's household expenditure survey, they find that adjustment for income quality effects reduces the urban–rural price differential by 7.7 percentage points (from 19.2 to 11.5 percent).

6.3. *Assessing the Plausibility of Our Estimates*

These low estimates of urban–rural price differentials are at odds with most of the existing literature for the Asia-Pacific region. For example, Ravallion and van de Walle (1991) find that the urban–rural price differential in Indonesia calculated over a basket consisting of food and housing is 10 percent, while Asra (1999), focusing on just food, finds it is 13–16 percent. Deaton and Dupriez (2011) obtain a differential of 11.5 percent for food prices in India, while Dikhanov (2010), focusing on food and clothing, finds it is 3.2 percent. Ravallion and Chen (2007), focusing on food and non-food consumption, obtain a differential for China of 19 percent in 1980, rising to 41 percent in 2002. Almås and Johnsen (2013), using Engel curves, obtain a differential for China of 65 percent in 1995, falling to 15 percent in 2002. Brandt and Holz (2006), and Gong and Meng (2008) compute spatial price differences across regions in China. While not explicitly discussing urban–rural price differentials, Brandt and Holz provide a table from which urban–rural price differentials can be calculated. From their Table 7 we obtain a price differential of 24 percent in 1990, rising to 31 percent or 40 percent in 2000 depending on the method used.

The explanation for this low difference in the urban–rural prices is probably that, due to cost considerations, the rural price quotes in ICP 2005 are not rural enough. In addition, the product lists in ICP 2005 were drawn up with urban consumers in mind (as is done in many cases in the consumer price index). It is therefore likely that quite a few products are representative in urban areas but unrepresentative in rural areas of the same country, while hardly any are representative in rural areas but not in urban areas. An analogy can be drawn here with Paasche and Laspeyres. An urban product list generates a Paasche-type index that underestimates the urban–rural price differential, while a rural product list generates a Laspeyres-type index that does the reverse. The Paasche analogy is applicable to ICP 2005.

Furthermore, Deaton and Dupriez (2011) show that even if rural price quotes are obtained, they may not reflect the actual prices faced by the rural consumers because a large portion of their consumption come from their own or neighbors' production. Dikhanov (2010) did not impute prices for own-produced goods which, as pointed out by Deaton and Dupriez, may be the reason he found such low urban–rural price differentials. ICP does not collect information on home produced goods and, therefore, understates the urban–rural price differentials.

This tendency of sampling a higher proportion of urban and unrepresentative products could cause the price levels in many Asia-Pacific countries to be overestimated and, consequently, the GDP in US dollars underestimated. The CPRD method is unable to deal with this situation since it does not allow the representativity of a product to vary within a country. Hence even when CPRUD is used, differences between urban and rural prices may be partially masked by the failure to account for the fact that often urban representative prices are being compared with rural unrepresentative prices.

6.4. *Implications for China in ICP 2005*

The average product prices for each country in ICP 2005 were, in most cases, calculated as a weighted average of the urban and rural price quotes, where the weights were supposed to reflect the relative urban and rural expenditure shares. This, however, is not true for China since all its prices are drawn from 11 cities. Hence all its price quotes are urban. To the extent that urban prices are higher than rural prices, sampling only from urban areas could have acted to push up the recorded price level in China relative to other countries. China's sampling only from urban areas could therefore partly explain why China emerged from ICP 2005 about 40 percent poorer than previously thought.

Our results can shed some light on this hypothesis. While we do not have any data for China itself, our measured urban–rural price differentials for the Asia-Pacific region are so small that it suggests either that the urban–rural distinction does not matter or more likely that the so-called "rural" price quotes from other countries are not rural enough. In other words, it seems that most of the countries in our sample also effectively only priced products in urban areas. Hence China's sampling only from urban areas should not have distorted much the results in ICP 2005.

A more likely candidate is excessive sampling in China of unrepresentative products rather than sampling from urban areas per se. Deaton and Heston (2010) make the point as follows:

> [T]he Chinese Bureau of Statistics chose the 11 cities because they were most likely to have outlets carrying the types of products and brands in the ICP specifications, and those prices are likely to be unrepresentatively high. (p. 21)

7. Conclusion

We have considered a number of problems with existing ICP methodology, their implications for international comparisons of real income, and some possible improvements and extensions that could be implemented in future rounds of ICP. Our analysis has been made possible by our unique dataset consisting of over 600,000 individual price quotes drawn from nine countries in the Asia-Pacific region. Being the first study of its kind, we undertook a comprehensive analysis of the problems and potential improvements in the construction of the ICP basic heading prices indexes. This is important since these basic heading price indexes are the fundamental building blocks on which the price comparisons at the national level are constructed.

In particular, we tried running CPD on the individual price quotes. The availability of individual price quotes also allowed us to estimate CPUD-type models that include urban dummies. While estimation of CPD and CPRUD-type models on the individual price quotes is not feasible in practice (due to confidentiality restrictions), it is still of interest to see how these methods perform empirically.

Going forward, we recommend that serious consideration be given in future rounds of ICP to WLS applied to the geometric means. Using WLS should provide some of the efficiency gains that would result from using the individual price quotes, while avoiding the serious drawbacks of this approach (i.e., confidentiality restrictions and distortions resulting from mismatches in the number of price quotes or in the share of urban and rural price quotes across countries). To do this, participating countries would need to supply the geometric means of the individual price quotes in each basic heading, along with the standard deviation of the logs of the price quotes instead of just the arithmetic mean as is currently the case. Providing this additional information would not violate confidentiality restrictions.

We support the ICP's decision to abandon the use of representative dummies. The underlying assumptions required to justify the use of representative dummies in a CPD-type model are quite stringent and were certainly not satisfied in our dataset. Nor are these conditions likely to be satisfied in future rounds of ICP. This is why importance-weighted CPD is being used in ICP 2011 instead of CPRD.

Another contribution of our study is that we show how CPD-type methods can be used to measure urban–rural price differentials across countries. This is an important topic in its own right that has implications for the poverty measurement literature. Our results also shed some light on the debate over the surprisingly low GDP estimate obtained for China in ICP 2005. One possible explanation is that this was the result of China only pricing products in urban areas. Our estimated urban–rural price differentials for the Asia-Pacific region are so small (i.e., 2.6 percent) as to suggest that either the urban-rural distinction is not empirically important, or more likely that the so-called "rural" price quotes in our sample are not rural enough. Either way the implication is that a lack of rural price quotes from China does not explain its lower than expected GDP as compared with other Asia-Pacific countries.

References

Almås, I. and A. A. Johnsen, "The Cost of Living in China: Implications for Inequality and Poverty," Memorandum 06/2013, Oslo University, Department of Economics, 2013.

Asra, A., "Urban–Rural Differences in Costs of Living and Their Impact on Poverty Measures," *Bulletin of Indonesian Economic Studies*, 35(3), 51–69, 1999.

Blades, D., "GDP and Main Expenditure Aggregates," chapter 3 in *ICP 2003–2006 Handbook*, The World Bank, Washington DC, 2007a.

———, *China in ICP 2005*, OECD, Paris, 2007b.

Brandt, L. and C. A. Holz, "Spatial Price Differences in China: Estimates and Implications," *Economic Development and Cultural Change*, 55(1), 43–86, 2006.

Chen, S. and M. Ravallion, "The Developing World is Poorer than We Thought, But No Less Successful in the Fight against Poverty," *Quarterly Journal of Economics*, 125(4), 1577–625, 2010.

Cuthbert, J. and M. Cuthbert, "On Aggregation Methods of Purchasing Power Parities," OECD Economics Department Working Papers, No. 56, OECD Publishing, Paris, 1988.

Deaton, A., "Measuring Poverty in a Growing World (Or Measuring Growth in a Poor World)," *Review of Economics and Statistics*, 87(1), 1–19, 2005.

Deaton, A. and O. Dupriez, "Using Unit-Values to Assess Spatial Price Differences: Evidence from India and Brazil," Paper presented at the ICP Technical Advisory Group Meeting in Washington DC, April 18–19, 2011.

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

Diewert, W. E., "Axiomatic and Economic Approaches to International Comparisons," in A. Heston and R. E. Lipsey (eds), *International and Interarea Comparisons of Income, Output, and Prices*, NBER Studies in Income and Wealth, Vol. 61, University of Chicago Press, Chicago and London, 13–87, 1999.

———, "On the Stochastic Approach to Linking the Regions in the ICP," Discussion Paper 04-16, Department of Economics, University of British Columbia, Vancouver, Canada, 2004.

———, "New Methodological Developments for the International Comparison Program," Discussion Paper 08-08, Department of Economics, University of British Columbia, Vancouver, Canada, 2008.

Dikhanov, Y., "Income Effect and Urban–Rural Price Differentials from the Household Survey Perspective," Paper presented at the 2nd ICP Technical Advisory Group Meeting, Washington DC, February 17–19, 2010.
Eltetö, O. and P. Köves, "On a Problem of Index Number Computation Relating to International Comparison," *Statisztikai Szemle*, 42, 507–18, 1964.
Feenstra, R. C., H. Ma, J. P. Neary, and D. S. P. Rao, "Who Shrunk China? Puzzles in the Measurement of Real GDP," *Economic Journal*, 123(12), 1100–29, 2013.
Gini, C., "On the Circular Test of Index Numbers," *International Review of Statistics*, 9(2), 3–25, 1931.
Gong, C. H. and X. Meng, "Regional Price Differences in Urban China 1986–2001: Estimation and Implications," Economics Program, Research School of Social Sciences, The Australian National University, Canberra, Australia, 2008.
Hill, R. J., "Measuring Substitution Bias in International Comparisons Based on Additive Purchasing Power Parity Methods," *European Economic Review*, 44, 145–62, 2000.
Hill, R. J. and T. P. Hill, "Recent Developments in the International Comparison of Prices and Real Output," *Macroeconomic Dynamics*, 13(Supplement no. 2), 194–217, 2009.
Hill, T. P., "Estimation of PPPs for Basic Headings Within Regions," chapter 11 in *ICP 2003–2006 Handbook*, The World Bank, Washington DC, 2007.
Kennedy, P. E., "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 71(4), 801, 1981.
Maddison, A., "Background Note on 'Historical Statistics' " (www.ggdc.net/Maddison), 2008.
Neary, J. P., "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income," *American Economic Review*, 94(5), 1411–28, 2004.
Rao, D. S. P., "The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power Parities in the ICP," Paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, Vancouver, Canada, June 30–July 3, 2004.
Ravallion, M. and S. Chen, "China's (Uneven) Progress Against Poverty," *Journal of Development Economics*, 82(1), 1–42, 2007.
Ravallion, M. and D. van de Walle, "Urban-Rural Cost-of-Living Differentials in a Developing Economy," *Journal of Urban Economics*, 29(1), 113–27, 1991.
Rouen, R. and C. Kai, "China's GDP in U.S. Dollars Based on Purchasing Power Parity," World Bank Policy Research Paper No. 1415, 1995.
Summers, R., "International Price Comparisons Based upon Incomplete Data," *Review of Income and Wealth*, 19(1), 1–16, 1973.
Szulc, B. J., "Indices for Multiregional Comparisons," *Przeglad Statystyczny* 3, *Statistical Review*, 3, 239–54, 1964.
Vogel, F., "ICP Survey Framework," Paper presented at the 3rd ICP Technical Advisory Group Meeting, Paris, France, June 10–11, 2010.
World Bank, *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program*, International Bank for Reconstruction and Development/World Bank, Washington DC, 2008.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix: Pooled Estimation of CPD-Type Models