Session Number: Parallel Session 7C
Time: Friday, August 27, AM

Paper Prepared for the 31st General Conference of
The International Association for Research in Income and Wealth

St. Gallen, Switzerland, August 22-28, 2010

The International Comparisons Program and the Country-Product-Dummy Method: An Application to the Asia-Pacific Region

Robert J. Hill and Iqbal Syed

For additional information please contact:

Name: Robert J. Hill
Affiliation: University of Graz, Austria

Email Address: robert.hill@uni-graz.at

The International Comparisons Program and the Country-Product-Dummy Method: An Application to the Asia-Pacific Region

Robert J. Hill[a] **and Iqbal Syed**[b]

21 July 2010

Abstract:
Basic heading (disaggregated) price indexes provide the building blocks from which international comparisons are made. Most errors arise at this disaggregated level and hence it is here that the most pressing research problems can be found. Perhaps the most striking result that emerged from the International Comparisons Program (ICP) results for 2005 was that China came out 40 percent smaller than previously thought. Using the raw price quote data from a sample of nine countries from the Asia-Pacific region we consider the extent to which this result can be explained by either an excessive focus in China on urban areas or on unrepresentative products in the data collection process. More generally, we consider whether the country-product-dummy (CPD) method used to construct the basic heading price indexes could be improved in future rounds of ICP by including representative dummies, correcting for heteroscedasticity and semilogarithmic coefficient bias, or by pooling the estimation of CPD equations across basic headings. We also explore the viability of estimating CPD-type regressions directly from the individual price quotes rather than country average prices as is currently done in ICP.
Keywords: International Comparisons Program; Country-Product-Dummy Model; Urban-Rural Price Differences; Representative and Unrepresentative Products; Product Specification
JEL Classification Codes: C43, O53.

[a]Department of Economics, University of Graz, Universitätsstrasse 15/F4, 8010 Graz, Austria: robert.hill@uni-graz.at
[b]School of Economics, University of New South Wales, Sydney 2052, Australia: i.syed@unsw.edu.au

* We thank the World Bank for funding this project and for providing the data, and Yuri Dikhanov for his comments.

1 Introduction

The latest round of the International Comparisons Program (henceforth ICP 2005) is a huge undertaking coordinated by the World Bank in collaboration with the OECD, Eurostat, IMF and UN. Its objective is to compare the purchasing power of currencies and real output across most countries in the world (146 countries participated in ICP 2005).

Perhaps the most surprising result that emerged from ICP 2005 was that China came out 40 percent smaller than previously thought (see Deaton and Heston 2008, Chen and Ravallion 2008 and Maddison 2008). This indicates that there may be a problem with either ICP 2005 or with earlier comparisons. ICP 2005 has the advantage that it is a much more detailed comparison and that for the first time China was an active participant. It should be noted though that China's participation in ICP 2005 was only on a limited scale, and the price quotes were obtained only from 11 cities and their surrounding areas (see Blades 2007a).

Deaton and Heston offer the following explanation for why there may be a problem with the ICP 2005 results for China:

> "Many of the qualities available in poorer countries are not available in higher income countries, while more of the qualities available of richer countries can also be found in poorer countries. ... The consequence is that prices for the ICP were often collected in higher-end outlets, which has the effect of raising price levels of poorer countries. This was made more likely in 2005 than previously because of the much closer review of prices across countries so that, for example, international brands were priced in (say) China, because they were available, even if mainly in high-end outlets. To the extent this happened, it would have the effect of raising parities in poorer countries, making them appear to have less income and output than in fact they do." (Deaton and Heston 2008)

In other words, this suggests that a higher proportion of the price quotes obtained for

China were unrepresentative as compared with richer countries.

If there is a problem in the ICP 2005 results for China, two possible culprits therefore could be that prices were collected only in urban areas or that the products priced were disproportionately unrepresentative.[1] While we do not have any data for China itself, we are able to partially address these issues using data from other countries in the Asia-Pacific region.

We return to these issues after some preliminary discussion of the mechanics of ICP 2005.

The ICP 2005 aggregate results at the level of GDP are obtained from 155 basic heading price indexes.[2,3] These basic heading price indexes 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 of the errors that occur, including those identified by Deaton and Heston, are likely to arise in the process of calculating the basic heading prices indexes. It is here at this disaggregated level that the most pressing research problems can be found.

Anecdotal evidence suggests that, other things equal, representative products tend to be cheaper than unrepresentative products (and that the same product is cheaper in a rural area than in an urban area). For this reason, Eurostat and the OECD have for many years asked countries to identify all priced products within each basic heading as either representative or unrepresentative in their internal comparisons, so that corrections for any imbalances can be made.

In ICP 2005 the Technical Advisory Group (TAG), of which both Deaton and Heston were members, recommended that comparisons at the basic heading level should

[1] These factors are not completely independent. Pricing only in urban areas in a developing country will tend also to imply pricing not enough representative products.

[2] Only 142 basic headings were used in the comparisons in the Asia-Pacific region.

[3] A basic heading is the lowest level of aggregation at which expenditure weights 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 2007b).

be made using an extended version of the country-product-dummy (CPD) model which also includes representative dummies. Hence all participating countries in ICP 2005 were asked to identify which of the products they priced were representative. However, this information was not actually used in the Asia-Pacific and most other regions.

One of the objectives of this study is to revisit this issue specifically for the Asia-Pacific region to see whether the decision to omit representative dummies was justified. In principle, the inclusion of representative dummies could correct for the type of bias described by Deaton and Heston. This, however, will only be the case if representative products are identified in a reasonably consistent way across countries.

Our data set consists of 605,998 price quotes drawn from 92 basic headings (covering most of household consumption) for nine countries in the Asia-Pacific region in 2005.[4]

Our findings are mixed. The inclusion of representative dummies undoubtedly increases the explanatory power of our CPD-type regressions. Most of the dummies are significant and have the expected sign. Hence the inclusion of representative dummies has the potential to at least partially alleviate the concerns of Deaton and Heston. However, at the same time it is clear that representative products were not identified in a consistent manner across countries. As a result, the inclusion of representative dummies could itself introduce noise and bias into the results.

Overall, we estimate the representative-unrepresentative price differential to be about 9 percent. While significant, and putting to one side the lack of consistency in the way that representative products were identified, this finding suggests that an excessive focus on unrepresentative products can explain only a relatively small portion of the downward revision for China.

We also consider the viability of further extending the basic CPD method to include urban and outlet-type dummies. Instead of including urban dummies, ICP 2005

[4]Strictly speaking we should refer to economies rather than countries, given that two of our sample (Hong Kong and Macao) are not countries. Nevertheless, we will henceforth use the term 'countries' since almost all the economies included in ICP 2005 are countries.

averaged prices within each country prior to application of the CPD method. One advantage of estimating the CPD model, inclusive of urban dummies, directly on the individual price quotes is that it is then possible to obtain estimates of the average price differential between urban and rural areas. We find that this differential is surprisingly small, about 3 percent.[5] In our opinion this result may reflect some problems with the treatment of rural and urban data in ICP 2005. More specifically, Deaton and Heston's argument that a higher proportion of products are unrepresentative in poorer countries may apply also within a country with more products being unrepresentative in rural areas than in urban areas.

Whether or not this is the case, assuming that the observed rural-urban price differential for China would have been similar to that of the other Asia-Pacific countries in our sample, our result suggests that the lack of rural data from China in ICP 2005 cannot explain much (in fact even less than the lack of representative products) of the large downward revision in the estimates of China's GDP. Hence we may have to look elsewhere for an explanation of the large downward revision in China's GDP arising from ICP 2005.

More generally, ICP 2005, and previous applications of CPD, have tended to neglect some important econometric issues. We find clear evidence of heteroscedasticity in the Asia-Pacific data set, and hence correct for it using feasible generalized least squares (FGLS). We also correct for semilogarithmic coefficient bias, which results from the fact that our basic heading price indexes are equal to the exponents of our estimated coefficients on the country dummies. We correct this bias using a version of Kennedy's (1981) formula. For most of our 92 basic headings the heteroscedasticity and semilogarithmic coefficient bias corrections are small. However, for a few headings it is quite large. These headings tend to be of the comparison-resistant variety and rely on only a small number of price quotes that often do not vary much within a country.

[5] Most studies for the Asia-Pacific region generally find rather larger rural-urban price differentials: see for example Ravallion and van de Walle (1991), Asra (1999), Brandt and Holz (2006), and Gong and Meng (2008).

We also consider whether simultaneous estimation of the CPD model over a group of basic headings in a seemingly unrelated regression (SUR) type setting can improve the efficiency of the estimated price indexes. Finally, as well as exploring the impact of each of these corrections on the basic heading price indexes, we also investigate the cumulative impact on the aggregate price indexes at the level of GDP. While our various corrections seem to have little impact on the aggregate price indexes themselves, we do find evidence of an increase in the observed level of relative price and relative quantity variability across countries.

A number of the issues raised here deserve closer scrutiny, and we hope some of our findings will be of use to future rounds of ICP.

2 The Country-Product-Dummy Method and its Extensions

The country-product-dummy (CPD) method, first proposed by Summers (1973), calculates the price index for a basic heading for all countries simultaneously.[6] The CPD model estimates the following regression equation:[7]

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

[6]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, for example, Eurostat and the OECD use the nonstochastic EKS-S method to construct their basic heading price indexes (see Hill and Hill 2009).

[7]It is common when estimating the CPD model to normalize the prices of both 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. The reason for our slightly nonstandard formulation of the CPD model will become apparent when we discuss the problem of semilogarithmic coefficient bias.

where p_{km} denotes the price of 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. 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. That is

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

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 ICP, p_{km} is an average of the price quotes obtained from all the outlets in country k. An alternative approach would be to include all the individual price quotes for product m directly in the CPD regression. We would then have multiple observations of p_{km}. In other words, p_{km} would be replaced by p_{kmr} where $r = 1, \ldots, R_k$ indexes the price quotes on product m available in country k. In the empirical comparisons later in the paper, this is the approach we use.

An extension of the CPD method, the country-product-representative-dummy (CPRD) method was proposed by Cuthbert and Cuthbert (1988). It simply adds an additional dummy variable to the model as follows:

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

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

The error term, $\hat{\varepsilon}_{km}$, for a product that is representative in country k should tend to be negative in the CPD model (since other things equal a representative product should be cheaper than an unrepresentative product). If representative products can be identified, this information can be utilized to correct for imbalances between the proportions of representative and unrepresentative price quotes within a basic heading across countries. In effect, either the prices of representative products can be adjusted upwards by a representativity factor or the prices of unrepresentative products can be adjusted downwards. The CPRD method estimates the adjustment factor simultaneously with the product and country factors.

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)

Unfortunately,

> "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. This means that some of the estimated basic heading price indexes in the Asia-Pacific region could be affected by the types of bias discussed by Deaton and Heston.

In principle, the CPRD method can be further extended, when the individual price quotes are available to include urban and outlet type dummies [i.e., the country-product-representative-product-urban-outlet-dummy (CPRUOD) method] as follows:

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

where w is a dummy that equals 1 if product m is from an urban area in country k and zero otherwise, while $i = 1, \ldots, I$ indexes a series of outlet types (e.g., 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. We assess the feasibility of using this extended model in an ICP context.

Econometrically the methods employed for estimating the CPD model in ICP could be improved. The CPD model in ICP 2005 is estimated using ordinary least squares (OLS). In the presence of heteroscedasticity the OLS standard errors will be biased. Our main focus here, however, is on the point estimates of the parameters of the model since the price indexes are derived directly from them. For this reason, our primary interest in heteroscedasticity is in its impact on the efficiency of our parameter estimates. In the presence of heteroscedasticity efficiency can be increased by using generalized least squares (GLS). At this point, it is sufficient to note that we find clear evidence of heteroscedasticity in our CPD-type regressions and hence there is a strong case for using GLS.

As noted above, the basic heading price indexes in CPD-type models are obtained from the exponents of the estimated β_k parameters. Goldberger (1968), under the assumption that the error term in the CPD-type regression equation is normal, shows that

$$E[\exp(\hat{\beta}_k)] = \exp\left[\beta_k - \frac{1}{2}\hat{\sigma}_k^2\right],$$

where $\hat{\sigma}_k^2$ is an estimate of the variance of $\hat{\beta}$. In other words, $\exp(\hat{\beta}_k)$ is a biased estimator of $\exp(\beta_k)$. To correct for this bias, Kennedy (1981) suggests the following estimator of $\exp(\beta_k)$, denoted here by $\tilde{p}_k = \widetilde{\exp}(\beta_k)$:

$$\widetilde{\exp}(\beta_k) = \exp\left[\hat{\beta}_k + \frac{1}{2}\hat{\sigma}_k^2\right]. \tag{2}$$

It is important when making this correction that none of the country price indexes are normalized to one. If the price index of country 1 is normalized to one, then by construction $\hat{\sigma}_1^2 = 0$ and hence the Kennedy correction reduces the price indexes of all countries except country 1. This will cause a violation of base country invariance. Given that the choice of base country is arbitrary, use of the Kennedy correction here will cause the price level in the base country to be systematically overestimated relative to all other countries.

It is for this reason that we specify a formulation of the CPD model in (1) that does not have a base country with price normalized to one. In this specification, the

Kennedy correction can be applied without creating any systematic biases in the price indexes. However, the results will now not be invariant to the choice of base product. One way to resolve this problem is to use each product in turn as the base, and then average the results. Given that the price indexes are relatively insensitive to the choice of base product, here we simply choose one as the base for each heading rather than using this averaging procedure.

Our last extension of the basic CPD-type model is to demonstrate how, rather than estimating a CPD-type model separately for each basic heading, we can pool headings in related categories and estimate the system of equations simultaneously. This has the potential to improve the efficiency of our parameter estimates, as well as allowing us to impose a common coefficient on the representative dummies, urban dummies or outlet-type dummies across groups of headings. Focusing on the case of the CPRUD model, letting $n = 1, \ldots, N$ index the basic headings included in the pool, the pooled version of the model is estimated as follows:

$$\ln p_{knm} = \sum_{n=1}^{N} \sum_{\mu=2}^{M_n} \alpha_{n\mu} x_{n\mu} + \sum_{n=1}^{N} \sum_{j=1}^{K} \beta_{jn} y_{jn} + \gamma z + \delta w + \sum_{i=2}^{I} \theta_i u_i + \varepsilon_{knm}, \tag{3}$$

Abstracting from the Kennedy correction, the country price indexes for each basic heading are obtained by exponentiating the estimated $\hat{\beta}_{kn}$ parameters:

$$\hat{p}_{kn} = \exp(\hat{\beta}_{kn}).$$

These can be compared across countries for the same basic heading (i.e., $\exp(\hat{\beta}_{kn} - \hat{\beta}_{jn})$) but should not be across basic headings for the same country (i.e., $\exp(\hat{\beta}_{kn_1} - \hat{\beta}_{kn_2})$) even when they are derived from the same CPD-type pooled regression. Comparisons of the latter type are not meaningful since there is no overlap in the product lists in two different basic headings. In an ICP context, comparisons of the first type are all that are needed from CPD-type methods. Aggregation across basic headings is done using standard price index formulas.

3 The Data Set

Our data set 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. In total there are 142 basic headings in ICP 2005 for the Asia-Pacific region. Our price quotes are drawn from 92 of these headings, all of which belong in the Final Consumption Expenditure by Households category.[8] Our list of basic headings is shown in Table 1.

Insert Table 1 Here

For our purposes the data set while large has some problems. Three countries (Fiji, Hong Kong and Malaysia) identified all products as representative, while Vietnam failed to identify 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. These are as follows: Department stores

Supermarkets

Open markets/stalls

Specialized shops (traditional outlets)

Wholesale and discount stores

[8]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 data set from 610,024 to 605,998.

Other stores

Some summary information is provided in Table 2.[9]

Insert Table 2 Here

4 CPD-Type Regression Results

4.1 Plausibility of the estimated representative, urban and outlet-type dummy variable coefficients

We consider first our most general CPD-type model. This may be referred to as the country-product-representative-urban-outlet-dummy (CPRUOD) model. We assume 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 equations are shown in Table 3. 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, representa-

[9]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 centre; 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.

tive 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 results broadly support this hypothesis, with 54 coefficients being positive (and 33 statistically significant) and only 26 being negative (with 11 statistically significant).[10]

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 stores categories, it is difficult to form any priors on them. The results are not really supportive of our priors. 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 counterintuitive for open markets and wholesale and discount stores. 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).

Insert Table 3 Here

We suspect that there may be serious inconsistencies with the ways that outlet types are identified across countries, and that this may explain the erratic results. We would recommend that in the next round 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 data set.

[10] 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.

For these reasons, we now exclude outlet-type dummies from our regression model. Hence our focus now is the country-product-representative-urban-dummy (CPRUD) model. The results are presented in Table 4. The sign of the representative coefficients here accords rather 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 Fiji, 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, 63 times positive (of which 45 are significant) and 21 times negative (of which 14 are significant). Also, shown in Table 4 are results for the CPRD method. The results for CPRD are similar to those obtained for the representative dummies in CPRUD.

Insert Table 4 Here

Given that out of the nine countries in our sample Fiji, Hong Kong and Malaysia identified all products as representative, while Vietnam left this column blank, it is far from clear that the inclusion of representative dummies would have improved the results in ICP 2005. In particular, the use of CPRD in this context would actually cause an upward bias in the resulting price indexes for Fiji, 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 CPRD for the Asia-Pacific region in ICP 2005. Nevertheless, at some point in the future (once countries identify representative products more consistently) the inclusion of representative dummies may be justified.

ICP 2005 already makes use of urban-rural identifiers in its calculation of country average prices prior to estimation of the CPD model. Our findings here suggest that estimation of a CPD-type model, inclusive of representative and urban dummies, directly from the individual price quotes is a viable alternative to the current practice based on average prices. We have serious doubts though whether the inclusion of outlet types, at least in the form available in ICP 2005, would improve the quality of the results.

4.2 Differences in estimated price indexes across methods

Our focus when comparing the results is on two issues. First, we assess the sensitivity of the results to the choice of method. Second, we check for evidence of systematic differences between the results generated by different methods. Taking the former first, the average change in the price indexes of each country as a result of switching from method x to method y is measured here as follows:

$$A_k(x,y) = \frac{1}{N}\sum_{n=1}^{N} \max(P_{kn}^x/P_{kn}^y, P_{kn}^y/P_{kn}^x),$$

where P_{kn} denotes the price index of country k for basic heading n (expressed as the number of units of currency that have the same purchasing power as 1 Hong Kong dollar). Also of interest is the maximum change in a basic heading price index, calculated as follows:

$$M_k(x,y) = \max_{n=1,\ldots,N}[\max(P_{kn}^x/P_{kn}^y, P_{kn}^y/P_{kn}^x)].$$

The average and maximum changes as measured by the A_k and M_k formulas are shown in Table 5 for the following pairs of methods:[11]

CPD-CPRD

CPD-CPRUD

CPRD-CPRUD

CPRD-CPRDhet

CPRDhet-CPRDhetken CPRUD-CPRUDhet

CPRUDhet-CPRUDhetken

For example, $A_k = 1.081$ for Bhutan in a comparison between CPD and CPRD. This means that the basic heading price indexes for Bhutan change on average by 8.1 percent as a result of switching from the CPD to CPRD model (although the direction of this change can differ from one heading to the next).

Insert Table 5 Here

[11]CPRDhet and CPRDhetken denote, respectively, CPRD corrected for heteroscedasticity and CPRD corrected for heteroscedasticity and incorporating Kennedy's correction of semilogarithm coefficient bias.

One must be careful comparing the A_k and M_k coefficients across countries for a few reasons. First, the results depend on the choice of base country (here Hong Kong). Second, 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 the more problematic headings. Third, often large values of $A_k(x, y)$ may be attributable primarily to differences in the underlying data sets 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 data set than the corresponding CPD results. Fourth, 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 country-product-urban-dummy (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 25.7 percent. As noted above, most of this change is probably attributable to the large differences in the data sets used to calculate the CPD and CPRD results, rather than inherent differences in the underlying methods.

The largest M_k coefficients in Table 5 are 3.35 observed in a comparison of CPD and CPRD for Fiji for basic heading 81 (Cultural services), and 3.33 and 3.55 observed in a comparison of CPD and CPRUD, respectively, for Fiji for heading 81 (Cultural services) and Sri Lanka for heading 86 (Accommodation services). In other words, the price index for Sri Lanka for the 'Accommodation services' basic heading changes by a factor of 3.55 as a result of including representative dummies. Again, most of these large differences are probably attributable to the small number of price quotes with representative-unrepresentative indicators available for this heading (only 30 out of 112 price quotes for Sri Lanka for heading 86 had representative-unrepresentative identifiers) and the large variations between these price quotes. The big differences, therefore, typically occur in difficult-to-measure headings such as 18=Other edible oils and fats, 20=Frozen, preserved or processed fruit and fruit-based products, 30=Spirits, 39=Maintenance and repair of the dwelling, 72=Telephone and telefax services, 81=Cultural services, 82=Newspapers, books and stationery, 86=Accommodation services.

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 could not be applied uniformly to 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. It remains to be seen whether the use of CPRUD and CPRUOD would be preferable to the current ICP methodology of constructing average prices for each heading by sampling from the available price quotes according to location (both in terms of whether it is urban or rural and the outlet type). In principle, though, it does seem likely that CPRD would be an improvement on CPD at least for some headings (as long as the representative-unrepresentative indicators are identified in a reasonably consistent manner across countries).

4.3 Differences in price level dispersion across methods

We now turn to the issue of whether there are systematic differences between the price levels derived from the CPD, CPRD and CPRUD methods. Price levels are obtained by dividing each price index by its corresponding average 2005 market exchange rate, with Hong Kong again normalized to 1. Systematic changes in price levels as a result of switching from CPD to CPRD could arise if for example a disproportionate share of the price quotes in say country k, relative to the others in our sample, are unrepresentative. The use of the CPRD method should in this case lower the measured relative price level in country k.

Rather than comparing all possible bilateral pairings of countries, here we simply consider whether the spread of the price levels across all nine countries rises or falls as a result of adopting the CPRUD method. Our measure of spread is given by the standard deviation of the logarithms of the price levels for each basic heading as follows:[12]

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

where P_{kn} again denotes the price index for basic heading n in country k, ER_k denotes the market exchange rate for country k, and $\overline{\ln(P_{kn}/ER_k)}$ is the average log price level for basic heading n.

We find that σ_n is higher for the CPRD method than for CPD for 47 headings and lower for 38 headings, as shown in Table 6.[13] To see whether the difference between the

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

[13]As was noted above, for 7 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.

CPD and CPRD σ_n coefficients is significant we use the normal approximation to the binomial distribution. Let X denote the number of basic headings for which the CPRD σ_n coefficient is larger than its corresponding CPD σ_n coefficient. X is approximately normally distributed with mean $N/2 = 42.5$ and variance $N/4 = 21.25$. A value of $X = 47$, implies a standard normal test statistic $Z = (X - 42.5)/\sqrt{21.25} = 0.976$, which is not significant at the 5 percent level. Hence we cannot reject the null hypothesis that there is no systematic difference between the price level dispersion coefficients of the CPD and CPRD methods.

Nevertheless, given that Fiji, Hong Kong and Malaysia identified all products as representative (and we assumed that Vietnam's price quotes were all representative), it follows that the price levels of these countries should tend to be higher relative to the other countries under CPRD than under CPD. We do indeed observe this pattern in the data for most headings (although not for all since representative-unrepresentative indicators in some countries are only available for a subset of price quotes and hence the underlying universe of price quotes over which CPD and CPRD price indexes are calculated are not exactly matched). This pattern, however, does not have any systematic impact on overall price dispersion since while Fiji, Hong Kong and Malaysia are three of the four highest priced countries in our sample (see Table 9), Vietnam is the country with the lowest price level. The inclusion of Vietnam in this group acts to prevent a noticeable increase in price level dispersion.

Insert Table 6 Here

The results from a comparison of CPD and CPRUD also shown in Table 6 are quite similar. The CPRUD price level dispersion σ_n is higher for 46 headings and lower for 39 headings. Using the normal approximation to the binomial, we obtain a test statistic of $Z = 0.759$ which is likewise not significant.

By contrast, in a comparison of CPRD with CPRUD, the CPRD σ_n coefficient is higher for 53 headings, and smaller for only 31 headings. In this case $Z = -2.400$ which is significant at the 5 percent level. This finding can be explained by the fact that all the price quotes from the three countries with highest overall price levels (see Table 9), namely Fiji, Hong Kong and Macao, are urban. The inclusion of urban dummies acts to lower the relative price levels in these three countries, thus reducing overall price level dispersion.

4.4 Correcting for heteroscedasticity

We test for heteroscedasticity in the CPRD and CPRUD models using the Breusch-Pagan (BP) test (see Breusch and Pagan 1979). The BP tests for our basic headings clearly reject the assumption of homoscedasticity. The BP F statistics are significant at the 1 percent level for most basic headings and at the 5 percent level for the remaining headings. Hence we reestimate the CPRD and CPRUD models using GLS. We calculate the GLS weights using a standard method. Let $\hat{e}_{kmr}$ denote the residual $p_{kmr} - \hat{p}_{kmr}$ on price quote r on product m in country k obtained from the estimated OLS model for a particular basic heading. We regress $\hat{e}^2_{kmr}$ on the explanatory variables of the model on the assumption that the variance of the OLS errors are functions of the explanatory variables. For the CPRUD models, the explanatory variables are country, product, representative and urban dummies. Let $\hat{g}$ denote the predicted values of the dependent variable obtained from the above regression, and in addition we define $\hat{h} = \exp(\hat{g})$. The weights are given by the reciprocals of the square root of $\hat{h}$. The variables are transformed by multiplying all the variables of the models by these weights. The feasible GLS (FGLS) estimates are obtained by applying OLS to the transformed variables. Given that our assumption that the variance of the OLS errors are functions of the explanatory variables is correct, as indicated by the BP tests, then our use of GLS should improve the efficiency of our estimated parameters, and hence also of our price indexes.[14] This rather than concern over possible bias in the standard errors is our

[14] For the case of CPD run on country average prices, Rao (2004) argues that these averages should be more reliable for those countries that have more price quotes. Assuming the price quotes are identically and independently distributed the implied heteroscedasticity of the country average prices can be modelled directly. However, we cannot use such an approach here since we estimate the CPD model directly from the individual price quotes.

primary concern with regard to heteroscedasticity.

One problem that can arise in the implementation of FGLS on the ICP data is that the estimated error $\hat{e}_{kmr}$ could be zero or very close to zero for one or more observations. We observe three different reasons why $\hat{e}_{kmr}$ could equal zero. First, in a few basic headings (e.g. 40=Water supply, 41=Electricity, 54=Pharmaceutical products, and 92=Other financial services n.e.c.) only a single price quote is available for one or more countries. Second, even if there are multiple price quotes from a country but these price quotes all relate to the same product and are all identical, then the estimated error on all these price quotes will be zero. This situation is observed for basic headings 61=Motor cycles and 68=Passenger transport by sea and inland waterways. Third, even if a country prices multiple products, but for one of these products it is the only country pricing it and all the price quotes on it are identical, then $\hat{e}_{kmr} = 0$ for these observations. Such cases are observed for 54=Pharmaceutical products, 59=Paramedical services and 92=Other financial services n.e.c. The best solution for this latter case is deletion of the product in question, since a minimum requirement for inclusion in the comparison is that a product should be priced by at least two distinct countries.

While zero estimated errors are easily identified, there may also be situations where the estimated error is close to zero. These observations may tend to get large weights under FGLS and may cause parameter instability in the resulting regression coefficients. It is to prevent such instability that in the first stage of FGLS we regress $\hat{e}^2_{kmr}$ instead of $\ln \hat{e}^2_{kmr}$, as is more usual (see Wooldridge 2003), on the explanatory variables. Then in the second stage set the weights are set equal to the reciprocal of the exponent of $\hat{g}$ as opposed to just the reciprocal of $\hat{g}$.[15]

The average and maximum changes as measured by the A_k and M_k coefficients from using GLS on the CPRD and CPRUD methods are shown in Table 5. The use of GLS has the biggest impact on basic headings 29 (Mineral waters, soft drinks, fruit and vegetable juices), 52 (Non-durable household goods), 65 (Passenger transport by

[15] We experimented also with setting $\hat{h} = 1 + \hat{g}$. The results were almost identical to those obtained with $\hat{h} = \exp(\hat{g})$.

railway), 66 (Passenger transport by road), and 86 (Accommodation services). The impact across countries of correcting for heteroscedasticity on the basic heading price indexes ranges on average from 0.5 percent and 2.1 percent for both the CPRD and CPRUD methods.

With regard to price level dispersion, GLS applied to the CPRD model generates larger σ_n coefficients than OLS for 32 basic headings, while for 60 headings we observe the opposite result (see Table 6). In this case $N = 92$ rather than 85 since for seven headings where we could not identify the representative effect we replace CPRUD with CPD. The test statistic obtained from the normal approximation to the binomial is $Z = -2.919$, which is significant at the 5 percent level. The results for CPRUD are similar. GLS generates large σ_n coefficients for 36 headings, and lower coefficients for 56 headings. Now $Z = -2.085$, which is again significant. It is not immediately obvious why correcting for heteroscedasticity should systematically reduce price level dispersion across the countries in our data set.

4.5 Correcting for semilog coefficient bias

The average and maximum changes as measured by the A_k and M_k coefficients from implementation of the Kennedy correction in (2) on the CPRD and CPRUD methods estimated using GLS are shown in Table 5. It can be seen that the average impact of the Kennedy correction is small. Its impact is biggest on Fiji for the basic heading 92 (Other services n.e.c.), where the correction changes the CPRD and CPRUD price index by 52 percent. The next highest change is 8 percent, which is observed for Indonesia 91 (Other financial services n.e.c). Basic headings that experience large Kennedy corrections imply that there are significant relative price differences across countries for the products in this heading. These price differences may be genuine, or they could signal the presence of poor quality data. Any heading that experiences a large Kennedy correction therefore should be closely scrutinized.

The Kennedy correction increases price level dispersion for both the CPRD and CPRUD methods (in both cases corrected for heteroscedasticity). The Kennedy corrected σ_n coefficients are larger for 66 and 65 out of 92 heading, respectively, for CPRD and CPRUD. The corresponding values of Z obtained from the normal approximation to the binomial are 4.170 and 3.962 both of which are highly significant. The combination of correcting for heteroscedasticity and semilogarithmic bias act to at least partially offset each other in terms of their impact on price level dispersion across the countries in our data set.

The finding that the Kennedy correction by itself seems to act to increase price level dispersion should probably not be taken too seriously given the negligible magnitude of the correction. In practice, for the vast majority of headings, the Kennedy correction is so small that it can be safely ignored.

4.6 Pooled estimation of CPD-type models

It is possible to divide the basic headings in Table 1 into groups of similar headings, and then estimate the CPD-type model for pools of headings as shown in (3) for the case of the CPRUD model. Following ICP 2005, here we sort the headings into 10 groups as shown in World Bank (2008, Appendix C). Pooling has the potential to improve the efficiency of the estimated basic heading price indexes, a point that has been raised in an ICP context recently by Silver (2009).

A number of caveats, however, apply. First, if a fully flexible model is estimated that allows all the estimated coefficients, including the representative and urban dummies to vary across basic headings, then pooling is equivalent to a seemingly unrelated regression (SUR) model (see Zellner 1962). Because there are no common variables across basic headings, however, the cross-equation correlations are zero and the estimated SUR coefficients collapse to the OLS coefficients. Consider, for example, the representative dummies. Though these dummies are common to all basic headings, the estimated coefficients differ across basic headings. In a SUR context, this means that the representative dummies across different basic headings are essentially different variables. The same holds for the urban and country dummy variables.

For pooling to have an impact it is necessary to impose restrictions on the coefficients across basic headings. These restrictions may take the form of equality constraints – such as the equality of the representative or urban dummies coefficients – across basic headings. The key issues are, first, whether the imposition of such restrictions is conceptually plausible, and, second, whether their imposition actually reduces the standard errors of the estimated coefficients. Conceptually, it is not clear whether such restrictions are desirable. Empirically, we find that out of eight groups, pooling of the CPRUD models with equality constraints increases the mean of the estimated standard errors in five groups (four of which are significant at the 5 percent level based on pair-wise Wilcoxon signed rank tests) and decreases the standard errors in three groups (only one of which is significant at the 5 percent level).[16] Similar results are obtained from a comparison of the CPRD pooled and un-pooled models. The fact that pooling with equality restrictions increases the estimated coefficient standard errors for five of the nine groups indicates that there are significant differences between the unconstrained representative and urban dummy coefficient estimates across basic headings. For example, in the food group, the estimated urban dummy coefficient ranges between -0.063 and 0.119 across basic headings with a mean of 0.032, while the estimated coefficient obtained from the pooled model is 0.035.

In summary, the case for pooling is at best mixed. It is something that might be worth considering for some groupings of basic headings in combination with equality restrictions on the representative and urban coefficients, particularly when a prior case can be made for imposing these restrictions. However, it should probably not be used on a regular basis.

Insert Table 7 Here

[16]Two groups, health and education, are excluded. This is because all the observations in the health category are representative and urban (since they are drawn only from Hong Kong and Macao), while for education we have only one basic heading.

5 Measuring Price Differences Between Urban and Rural Areas and Between Representative and Unrepresentative Products

The CPRUD regression model takes the following form:

$$\ln p_{km} = \kappa + \sum_{\mu=2}^{M} \alpha_\mu x_\mu + \sum_{j=2}^{K} \beta_j y_j + \gamma z + \delta w + \varepsilon_{km},$$

where γ and δ denote, respectively, the coefficients on the representative and urban dummies. Estimating the CPRUD model for each basic heading n, we obtain 92 estimated coefficients $\hat{\gamma}_n$ and $\hat{\delta}_n$.

Focusing first on representativity, abstracting from semilog coefficient bias, the term $\exp(\hat{\gamma}_n)$ can be interpreted as a price index measuring the average price difference between representative and unrepresentative products, other things equal, for basic heading n. Given that unrepresentative products are the numeraire in our CPRUD formulation, and unrepresentative products tends to be more expensive than representative products, it follows that $\hat{\gamma}_n$ should be negative, and hence $\exp(\hat{\gamma}_n)$ less than one. An aggregate price index over all basic headings, with unrepresentative products as the numeraire, can be calculated in two stages. First we average across basic headings for each country k:

$$P^k_{Un,Rep} = \prod_{n=1}^{N} \left\{ [\exp(\hat{\gamma}_n)]^{s^k_n} \right\},$$

where s^k_n denotes the expenditure share of basic heading n in country k.[17] The overall price index is then obtained by averaging across countries as follows:

$$P_{Un,Rep} = \prod_{k=1}^{K} \left[\left(P^k_{Un,Rep} \right)^{s_k} \right], \tag{4}$$

where s^k denotes country k's share of total GDP. These weights are calculated as follows:

$$s^k = \frac{GDP^k/P^k}{\sum_{j=1}^{K} GDP^j/P^j},$$

[17] The reason for taking a geometric mean rather than say an arithmetic mean is so that the resulting price indexes do not depend on whether unrepresentative or representative products are chosen as the numeraire.

where GDP_k denotes GDP in country k denominated in units of own currency and P^k denotes the official ICP purchasing power parity (or price index) for country k. The ratios GDP^k/P^k generate GDPs for the k countries all denominated in units of the base country's currency. The term $P_{Un,Rep}$ is a price index for representative products Rep with the price of unrepresentative products Un normalized to 1. For example $P_{Un,Rep} = 0.8$ would imply that the prices of unrepresentative products on average are 25 percent higher than the prices of representative products.

Corresponding urban-rural price indexes are calculated in an analogous manner. The term $\exp(\hat{\delta}_n)$ can be interpreted as a price index measuring the average price difference between urban and rural products, other things equal, for basic heading n, with rural products as the numeraire (hence we expect that $\exp(\hat{\delta}_n) > 1$.

$$P^k_{Rur,Urb} = \prod_{n=1}^{N} \left\{ \left[\exp(\hat{\delta}_n) \right]^{s^k_n} \right\},$$

$$P_{Rur,Urb} = \prod_{k=1}^{K} \left[\left(P^k_{Rur,Urb} \right)^{s_k} \right], \tag{5}$$

with the weights calculated in exactly the same way.

In Table 8 we present the $P^k_{Un,Rep}$, $P_{Un,Rep}$, $P^k_{Rur,Urb}$ and $P_{Rur,Urb}$ price indexes. Overall we find that unrepresentative products on average are 9.0 percent more expensive that representative products [since $1/0.918 = 1.090$] while products on average are 3.2 percent more expensive in urban areas than in rural areas.

Insert Table 8 Here

The estimates of $P_{Un,Rep}$ and $P_{Rur,Urb}$ in Table 8 are perhaps lower than one might expect, particularly the latter.

It might be possible to explain the low value of $P_{Rur,Urb}$ using a variant on Deaton and Heston's argument. Deaton and Heston argue that a higher proportion of the products priced in poorer countries are unrepresentative. Similarly, it may be the case that a higher proportion of the products priced in rural areas are unrepresentative. In other words, it may be necessary to extend the concept of representativity. At present, users of this concept (such as the OECD and Eurostat) assume that a product is either representative or not in a country. However, it may be the case that within the same country it is representative in urban areas but not in rural areas. This effect may be particularly applicable to poorer countries. If so, we should expect the $P_{Rur,Urb}$ price differential estimate for our sample of Asia-Pacific countries to be too low. Furthermore, if this effect holds true, then poorer countries with the largest rural population shares will tend to have their price levels overestimated and hence per capita incomes underestimated.

6 Correcting for Differences in the Price Quote and Urban-Rural Expenditure Mixes Across Countries

Hong Kong is 100 percent urban both in terms of its price quotes and population. The CPRUD method will 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 is not justified since households in Hong Kong do not have the option of purchasing in rural areas (without travelling beyond its borders). The problem here is that the CPRUD method implicitly assumes that the expenditure mix across urban and rural areas is the same in all countries, which it is not. Hence to prevent bias an adjustment is required. The appropriate scale factors can be derived from the CPRUD urban-rural price indexes in Table 8. Let Exp^k_{Urb} and Exp^k denote urban and total expenditure, respectively, in country k. One possible way of adjusting the CPRUD basic heading price indexes is as follows:

$$\tilde{P}^k_n = \left[\left(\frac{Exp^k_{Urb}}{Exp^k} \right) P_{Rur,Urb} + 1 \right] P^k_n, \tag{6}$$

where P^k_n denotes the original CPRUD price index for basic heading n in country k, $\tilde{P}^k_n$ is the adjusted index, and $P_{Rur,Urb}$ is the urban-rural price index derived from (5).[18] From (6) we can see for a totally urban population such as Hong Kong that

[18]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.

$\tilde{P}_n^k = (P_{Rur,Urb} + 1)P_n^k$, while for a totally rural population $\tilde{P}_n^k = P_n^k$. In other words, the more urban is total expenditure the bigger the upward adjustment in the price index and corresponding price level. Also, 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 they are invariant to rescaling). That is, in this case the CPRUD method gives the right answer.

Our conclusions here should be treated as preliminary. For example, it might be better to use basic heading specific urban-rural price indexes $P_{Rur,Urb,n}$ in (6) rather than the same price index for all headings. Also, we have not actually calculated numerical estimates of the adjustment factor in (6) for any of the countries in our data set. This whole topic of urban-rural adjustment factors for the CPRUD method warrants further investigation.

Is a similar adjustment required for representativity for the CPRD or CPRUD methods? In our opinion the answer is not necessarily. The concept of representativity is somewhat vague and is likely to be interpreted in different ways by different countries unless they are given very precise guidelines. For it to be useful, it is critical that countries use the same definition. One possible definition is as follows: a representative product in country k is one of the top 50 percent of products bought there (weighted by expenditure) in that particular basic heading.[19] Our example, helps illustrate the key difference between representative and urban indicators. It is possible for 99 percent of expenditure in country k to be urban, but it is not possible for 99 percent of expenditure to be on representative products.[20]

[19] Here we abstract from the issue mentioned above that a particular product may be representative in urban areas but not rural areas of the same country.

[20] One potential source of confusion over the concept of representativity is that some basic headings themselves are inherently more representative than others in each country. For example, the headings spirits, wines and beers could all three, along with all the products within each of these headings, be deemed unrepresentative in a predominantly Muslim country such as Indonesia. Representativity, in a CPD context, however is really a relative concept. Focusing on the beer example above, Indonesia should identify those beers that are most representative, rather than simply classify them all as

It does seem likely that expenditure in poorer countries is concentrated on a smaller range of products. If so, it follows that the proportion of representative products in the ICP product list will tend to be lower for poorer countries, and hence that the CPD method will tend to systematically underestimate price differences (and overestimate income differences) across countries. This is exactly the effect described by Deaton and Heston (2008). Methods such as CPRD and CPRUD, however, will only help to offset this bias if representative products are identified in a consistent way across countries (which does not seem to have been the case in ICP 2005 at least in the Asia-Pacific region).

7 Results at the Level of GDP

The overall ICP 2005 comparisons in the Asia-Pacific region cover 142 basic headings and 23 countries. We have recalculated the price indexes for 92 of these basic headings for nine countries (although for many headings our coverage of these nine countries is incomplete). We now consider the impact on the overall results at the level of GDP of replacing the official ICP basic heading price indexes calculated using the CPD method with our price indexes where available calculated using the CPRUD method corrected for heteroscedasticity and semilogarithmic coefficient bias.

The official ICP 2005 aggregate indexes for the Asia-Pacific region were calculated using the Gini-Eltetö-Köves-Szulc (GEKS) method (see Gini 1931, Eltetö and Köves 1964 and Szulc 1964). The GEKS method calculates the aggregate price index for country k as follows:

$$P_k = \prod_{j=1}^{K} \left[\left(\frac{\sum_{n=1}^{N} p_{kn} q_{jn}}{\sum_{n=1}^{N} p_{jn} q_{jn}} \right) \left(\frac{\sum_{n=1}^{N} p_{kn} q_{kn}}{\sum_{n=1}^{N} p_{jn} q_{kn}} \right) \right]^{1/(2K)} = \prod_{j=1}^{K} \left(P_{jk}^{P} \times P_{jk}^{L} \right)^{1/(2K)},$$

where p_{kn} denotes the price index of basic heading n in country k, q_{kn} denotes the corresponding quantity obtained by deflating expenditure on basic heading n by the price index p_{kn}, and P_{jk}^{P} and P_{jk}^{L} denote Paasche and Laspeyres price indexes, respectively.

unrepresentative.

Letting b denote the base country (here Hong Kong), the price indexes are typically rescaled so that the price index in the base country equals 1 (this is achieved by dividing the price index of each country P_k by the price index of the base country P_b). The official ICP 2005 basic heading aggregate price indexes and price levels, and our revised estimates are shown in Table 9.

Insert Table 9 Here

The σ_n coefficients at the aggregate level for the official ICP results and our revised estimates are quite similar at 0.369 and 0.385 respectively. In total, as a result of gaps in our data, close to half of the basic heading price indexes in our revised estimates data set are the same as in the official ICP results. If we were able to fill these gaps, the difference between the two σ coefficients might be larger.

Also of interest are the Paasche-Laspeyres spreads (PLS) between pairs of countries, defined here as follows:[21]

$$PLS_{jk} = \max(P^L_{jk}/P^P_{jk}, P^P_{jk}/P^L_{jk}).$$

The PLS may be interpreted as a measure of the extent of variability in relative prices p_{kn}/p_{jn} and quantities q_{kn}/q_{jn} across basic headings n for a pair of countries j and k. That is, a higher PLS implies greater variability in relative prices and quantities.

Matrices of Paasche-Laspeyres spreads defined on our set of nine countries derived from the official ICP data and our revised estimates are shown in Table 10. There are a total of 36 distinct bilateral comparisons that can be made between pairs of countries in our data set (i.e., $K(K-1)/2$ where $K = 9$). For 30 of these 36 bilateral comparisons, our PLS are larger than those obtained from the official ICP 2005 basic heading data. This suggests that there is greater variability in the price and quantity vectors across countries in our revised basic headings, and hence that the inclusion of representative

[21] See Hill (1999) for a discussion of the properties of the PLS. In particular, by construction $PLS_{jk} \geq 1$, and $PLS_{jk} = PLS_{kj}$. When either the price vectors or baskets of goods and services differ across countries j and k by only a scalar multiple then $PLS_{jk} = 1$ and there is no index number problem since all price index formulas should give the same answer. In such cases the data are consistent with the conditions for Hicks and Leontief aggregation respectively.

and urban dummies in the regression equation may enable us to better discern price differences across countries, which may otherwise be masked by mismatching of price quotes across countries.[22]

Insert Table 10 Here

8 Conclusion

We have considered a number of possible ways in which the ICP methodology could be extended in future rounds. First, there is the issue of whether the CPD-type method should include representative dummies. Given that out of the nine countries in our sample Fiji, Hong Kong and Malaysia identified all products as representative, while Vietnam left this column blank, it is far from clear that the inclusion of representative dummies we support the decision to exclude representative dummies in ICP 2005. Nevertheless, we think that at some point in the future (once countries identify representative products more consistently) the inclusion of representative dummies may be justified, although correction factors may then be required to prevent bias. We have shown one way in which these correction factors could be calculated.

With regard to the identification of location of purchases (i.e., urban or rural and outlet type), the question is not so much whether urban and outlet type dummies should be included in a CPD-type model as whether country average prices should be calculated or not prior to application of a CPD-type method. ICP 2005 computes country average prices that weight the individual price quotes depending on the location (i.e., urban or rural) of the purchase. The alternative is to apply a CPD-type method, inclusive of representative, urban and outlet-type dummies, directly to the individual price quotes. While in theory we prefer the latter approach, the lack of consistency in the location information across countries in ICP 2005 makes the former approach more appealing given the current state of the data.

[22]The expenditure data are the same in both comparisons. So changes in relative quantities here arise implicitly from changes in relative prices.

A strong case can be made on econometric grounds for correcting for heteroscedasticity and semilogarithmic coefficient bias in CPD-type regressions. In practice, however, the impact of these corrections is generally quite small. Pooling of CPD-type models during estimation as a means of increasing efficiency is an issue that perhaps deserves further attention. Given our preliminary analysis of this topic, we do not recommend doing this as a general rule. It may, however, be worth considering for certain groups of headings.

Finally, we have shown how CPD-type models can be used to quantify the price differential between representative and unrepresentative products, and between urban and rural locations. For our data set, we find that prices in urban areas are about 3.2 percent higher than in rural areas, while unrepresentative products are about 9 percent more expensive than representative products. We suspect that our estimated rural-urban price differential may be too low, perhaps due to a general tendency for products to be more representative in urban areas than rural areas, particularly in poorer countries.

Whether or not this is the case, our results have a direct bearing on the debate over the causes of the substantial downward revision in China's GDP arising out of ICP 2005. They suggest that the majority of this revision cannot be attributed to either an excessive sampling of unrepresentative products or of outlets in urban areas in China in ICP 2005. To be clear, we are not saying that the price differential between rural and urban areas is not significant, but rather that the price quotes and methods used in ICP 2005 are not able to fully capture this difference. Similarly, regarding the representative-unrepresentative price differentials, a note of caution is also required. The lack of consistency in the way representative products were identified across countries could have caused this differential to be seriously mismeasured.

In conclusion, we have raised a number of issues here that we think should be investigated further, and that may be of interest to future rounds of ICP, and more generally to anyone interested in comparing income levels and prices across countries.

References

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

Blades, Derek (2007a), "China in ICP 2005", Mimeo.

Blades, Derek (2007b) GDP and Main Expenditure Aggregates. Chapter 3 in *ICP 2003-2006 Handbook*. Washington DC: The World Bank.

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

Breusch, Trevor and Adrian Pagan (1979), "A Simple Test for Heteroscedasticity and Random Coefficient Variation", *Econometrica* 50, 987-1007.

Chen, Shaohua and Martin Ravallion (2008), "China is Poorer than we Thought, But No Less Successful in the Fight against Poverty", Policy Research Working Paper 4621. Washington DC: The World Bank, Development Research Group.

Cuthbert, Jim and Margaret Cuthbert (1988) On Aggregation Methods of Purchasing Power Parities. *OECD Economics Department Working Papers*, No. 56. Paris: OECD Publishing.

Deaton, Angus and Alan Heston (2008), "Understanding PPPs and PPP-Based National Accounts", NBER Working Paper No. 14499.

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

Eltetö, Oded and Pal Köves (1964), "On a Problem of Index Number Computation Relating to International Comparison", *Statisztikai Szemle* 42, 507-518.

Gini, Corrado (1931), "On the Circular Test of Index Numbers", *International Review of Statistics*, Vol.9, No.2, 3-25.

Goldberger, Arthur S. (1968), "The Interpretation and Estimation of Cobb-Douglas Functions", *Econometrica* 35, 464-472.

Gong, Cathy H. and Xin Meng (2008), "Regional Price Differences in Urban China 1986-2001: Estimation and Implications", Economics Program, Research School of Social Sciences, The Australian National University, Canberra, Australia.

Hill, Robert J. (1999), "Comparing Price Levels Across Countries using Minimum Spanning Trees", *Review of Economics and Statistics* 81(1), February 1999, 135-142.

Hill, Robert J. and T. Peter Hill (2009), "Recent Developments in the International Comparison of Prices and Real Output", *Macroeconomic Dynamics* 13 (supplement no. 2), 194-217.

Hill, T. Peter (2007), Estimation of PPPs for Basic Headings Within Regions. Chapter 11 in *ICP 2003-2006 Handbook*, Washington, DC: The World Bank.

Kennedy, Peter E. (1981), "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations", *American Economic Review* 71 (4), 801.

Maddison, Angus (2008), Background Note on "Historical Statistics," ≪www.ggdc.net/Maddison≫.

Rao, D. S. Prasada (2004), "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, 30 June-3 July, 2004.

Ravallion Martin and Dominique van de Walle (1991), "Urban-Rural Cost-of-Living Differentials in a Developing Economy", *Journal of Urban Economics* 29(1), 113-127.

Silver, Mick (2009), "The Hedonic Country Product Dummy Method and Quality Adjustments for Purchasing Power Parity Calculations", International Monetary Fund Working Paper WP/09/271. International Monetary Fund: Washington DC.

Summers, Robert (1973), "International Price Comparisons Based upon Incomplete Data", *Review of Income and Wealth* 19(1), 1-16.

Szulc, Bohdan J. (1964), "Indices for Multiregional Comparisons", *Przeglad Statystyczny* 3, *Statistical Review* 3, 239-254.

Wooldridge, Jeffrey (2004), *Introductory Econometrics: A Modern Approach.* Thomson South Western: Ohio, Second Edition.

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

Zellner, Arnold (1962), "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias", *Journal of the American Statistical Association* 57, 348-368.

Table 1: Our List of ICP Basic Headings for Final Consumption Expenditure by Households

1	**110111.1**	**Rice**
2	**110111.2**	**Other cereals, flour and other cereal products**
3	**110111.3**	**Bread**
4	**110111.4**	**Other bakery products**
5	**110111.5**	**Pasta products**
6	**110112.1**	**Beef and Veal**
7	**110112.2**	**Pork**
8	**110112.3**	**Lamb, mutton and goat**
9	**110112.4**	**Poultry**
10	**110112.5**	**Other meats and meat preparations**
11	**110113.1**	**Fresh, chilled or frozen fish and seafood**
12	**110113.2**	**Preserved or processed fish and seafood**
13	**110114.1**	**Fresh milk**
14	**110114.2**	**Preserved milk and other milk products**
15	**110114.3**	**Cheese**
16	**110114.4**	**Eggs and egg-based products**
17	**110115.1**	**Butter and Margarine**
18	**110115.3**	**Other edible oils and fats**
19	**110116.1**	**Fresh or chilled fruit**
20	**110116.2**	**Frozen, preserved or processed fruit and fruit-based products**
21	**110117.1**	**Fresh or chilled vegetables other than potatoes**
22	**110117.2**	**Fresh or chilled potatoes**
23	**110117.3**	**Frozen, preserved or processed vegetables and vegetable-based products**
24	**110118.1**	**Sugar**
25	**110118.2**	**Jams, marmalades and honey**
26	**110118.3**	**Confectionery, chocolate and ice cream**
27	**110119**	**Food products n.e.c.**
28	**110121**	**Coffee, tea and cocoa**
29	**110122**	**Mineral waters, soft drinks, fruit and vegetable juices**
30	**110211**	**Spirits**
31	**110212**	**Wine**
32	**110213**	**Beer**
33	**110220**	**Tobacco**
34	**110311**	**Clothing materials, other articles of clothing and clothing accessories**
35	**110312**	**Garments**
36	**110314**	**Cleaning, repair and hire of clothing**
37	**110321**	**Shoes and other footwear**
38	**110322**	**Repair and hire of footwear**
39	**110430**	**Maintenance and repair of the dwelling**
40	**110441**	**Water supply**
41	**110451**	**Electricity**
42	**110452**	**Gas**
43	**110453**	**Other fuels**
44	**110511**	**Furniture and furnishings**
45	**110512**	**Carpets and other floor coverings**

46	**110520**	**Household textiles**
47	**110531**	**Major household appliances whether electric or not**
48	**110532**	**Small electric household appliances**
49	**110533**	**Repair of household appliances**
50	**110540**	**Glassware, tableware and household utensils**
51	**110552**	**Small tools and miscellaneous accessories**
52	**110561**	**Non-durable household goods**
53	**110562.1**	**Domestic services**
54	**110611**	**Pharmaceutical products**
55	**110612**	**Other medical products**
56	**110613**	**Therapeutical appliances and equipment**
57	**110621**	**Medical Services**
58	**110622**	**Dental services**
59	**110623**	**Paramedical services**
60	**110711**	**Motor cars**
61	**110712**	**Motor cycles**
62	**110713**	**Bicycles**
63	**110722**	**Fuels and lubricants for personal transport equipment**
64	**110723**	**Maintenance and repair of personal transport equipment**
65	**110731**	**Passenger transport by railway**
66	**110732**	**Passenger transport by road**
67	**110733**	**Passenger transport by air**
68	**110734**	**Passenger transport by sea and inland waterway**
69	**110736**	**Other purchased transport services**
70	**110810**	**Postal services**
71	**110820**	**Telephone and telefax equipment**
72	**110830**	**Telephone and telefax services**
73	**110911**	**Audio-visual, photographic and information processing equipment**
74	**110914**	**Recording media**
75	**110915**	**Repair of audio-visual, photographic and information processing equipment**
76	**110921**	**Major durables for outdoor and indoor recreation**
77	**110931**	**Other recreational items and equipment**
78	**110933**	**Gardens and pets**
79	**110935**	**Veterinary and other services for pets**
80	**110941**	**Recreational and sporting services**
81	**110942**	**Cultural services**
82	**110950**	**Newspapers, books and stationery**
83	**110960**	**Package holidays**
84	**111000**	**Education**
85	**111110**	**Catering services**
86	**111120**	**Accommodation services**
87	**111211**	**Hairdressing salons and personal grooming establishments**
88	**111212**	**Appliances, articles and products for personal care**
89	**111231**	**Jewellery, clocks and watches**
90	**111232**	**Other personal effects**
91	**111262**	**Other financial services n.e.c**
92	**111270**	**Other services n.e.c.**

Table 2: Some Summary Information on Each Country

Countries	Outlet type	Region (Rural/Urban)	Representative (Y/N)	Number of Price Quotes
Bhutan	Yes	No	Yes	17085
Fiji*	Yes	Yes	Yes	9897
Hong Kong	Yes	No	Yes	45231
Indonesia	No	No	Yes	62972
Macao	Yes	No	Yes	28554
Malaysia	Yes	Yes	Yes	70683
Philippines	Yes	Yes	Yes	142379
Sri Lanka	No	Yes	Yes	72562
Vietnam	No	Yes	No	156635
TOTAL				605998

* Many of the price quotes do not have outlet type and representativity identifiers. Also all price quotes are identified as urban.

Table 3: Some Statistics on the Signs and Significance Levels of the Estimated Coefficients of the CPRUOD Model

Variables	Statistics	All coefficients	Positive	Negative
Representative variable				
	Number of +ve/-ve sign		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

*The base outlet type is Supermarkets

Table 4: Some Statistics on the Signs and Significance Levels of the Estimated Coefficients of the CPRD and CPRUD Models

Model	Variable/Statistics	All coefficients	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

Table 5: Average and Maximum Changes in Price Indexes by Method

A_k	CPD CPRD	CPD CPRUD	CPRD CPRUD	CPRD CPRDhet	CPRDhet CPRDhetken	CPRUD CPRUDhet	CPRUDhet CPRUDhetken
BHU	1.0814	1.0799	1.0057	1.0082	1.0020	1.0082	1.0020
FIJ	1.2570	1.2563	1.0041	1.0133	1.0129	1.0138	1.0130
INO	1.0195	1.0328	1.0234	1.0069	1.0032	1.0075	1.0033
MAC	1.0112	1.0108	1.0021	1.0054	1.0003	1.0054	1.0003
MAL	1.0106	1.0123	1.0053	1.0072	1.0002	1.0075	1.0002
PHI	1.0312	1.0324	1.0061	1.0108	1.0003	1.0110	1.0003
SRI	1.0829	1.0857	1.0157	1.0214	1.0007	1.0208	1.0008
VIE	1.0107	1.0211	1.0161	1.0086	1.0002	1.0090	1.0002

M_k	CPD CPRD	CPD CPRUD	CPRD CPRUD	CPRD CPRDhet	CPRDhet CPRDhetken	CPRUD CPRUDhet	CPRUDhet CPRUDhetken
BHU	1.3444	1.8141	1.0601	1.0722	1.0373	1.0678	1.0378
FIJ	3.3480	3.3293	1.0326	1.4871	1.5210	1.5104	1.5231
INO	1.0752	1.2070	1.0808	1.0459	1.0839	1.0530	1.0875
MAC	1.0946	1.1357	1.0207	1.0833	1.0035	1.0820	1.0035
MAL	1.0751	1.0787	1.0276	1.0682	1.0037	1.0701	1.0043
PHI	1.2268	1.2793	1.0297	1.2681	1.0055	1.2632	1.0062
SRI	1.1422	3.5485	1.0773	2.2010	1.0073	2.1315	1.0084
VIE	1.0635	1.1014	1.0597	1.1620	1.0046	1.1687	1.0058

Worst performing basic heading

	CPD CPRD	CPD CPRUD	CPRD CPRUD	CPRD CPRDhet	CPRDhet CPRDhetken	CPRUD CPRUDhet	CPRUDhet CPRUDhetken
BHU	82	82	86	66	81	66	81
FIJ	81	81	92	52	92	52	92
INO	18	18	19	29	91	29	91
MAC	20	20	76	86	58	86	58
MAL	72	86	19	65	65	65	65
PHI	30	30	30	52	65	52	65
SRI	86	86	92	86	65	86	65
VIE	30	30	19	66	65	66	65

Table 6: A Comparison of Price Level Dispersion Across Methods

x	CPD	CPD	CPRD	CPRD	CPRDhet	CPRUD	CPRUDhet
y	CPRD	CPRUD	CPRUD	CPRDhet	CPRDhetken	CPRUDhet	CPRUDhetken
$\sigma x > \sigma y$	38	39	55	60	26	56	27
$\sigma x < \sigma y$	47	46	29	32	66	36	65
Z	-0.976	-0.759	2.837	2.919	-4.170	2.085	-3.962

Table 7: Categories for Pooled Estimation of CPD-Type Models

		Number of basic headings	
		ICP 2005	Our data set
1	Food and non-alcoholic beverages	29	29
2	Alcohol, tobacco	5	4
3	Clothing and footwear	5	5
4	Housing, water, electricity, gas and other fuels	7	5
5	Furnishings, household equipments and maintenance	13	10
6	Health	7	6
7	Transport	13	10
8	Communication, recreation and culture	16	14
9	Education	1	1
10	Restaurants, hotels, miscellaneous goods and services	12	8
	TOTAL	108	92

Note: our list of basic headings here is restricted to those belonging to Final Consumption Expenditure by Households.

Table 8: Average Effects of Representative and Urban Indicators

Country	Rep Dummy	Urban Dummy	$P_{Un,Rep}$	$P_{Rur,Urb}$
Bhutan	-0,11102	0,02535	0,8949	1,0257
Fiji	-0,11666	0,03637	0,8899	1,0370
Hongkong	-0,05941	0,03392	0,9423	1,0345
Indonesia	-0,09923	0,03444	0,9055	1,0350
Macao	-0,05301	0,03564	0,9484	1,0363
Malaysia	-0,11843	0,02453	0,8883	1,0248
Philippines	-0,06154	0,02849	0,9403	1,0289
Srilanka	-0,04565	0,0301	0,9554	1,0306
Vietnam	-0,06765	0,03516	0,9346	1,0358
Average Effect (weighted)	-0,08602	0,0317	0,9176	1,0322

Note: The country rep-dummy and urban-dummy coefficients are the weighted averages of the corresponding estimated dummy coefficients at the basic heading level, where weights are given by the expenditure shares of the basic headings for a particular country. The country averages differ because the weights corresponding to the basic headings differ between countries. The average of all countries is the weighted average of all the country averages, where weights correspond to relative GDP (PPP adjusted) of the respective countries.

Table 9: GEKS Price Indexes and Price Levels for GDP

	Price Index Official data	Price Index Revised data	Price Level Official data	Price Level Revised data
Bhutan	2.7471	2.5439	0.4845	0.4486
Fiji	0.2479	0.2539	1.1403	1.1678
Hong Kong	1.0000	1.0000	1.0000	1.0000
Indonesia	686.4862	626.9158	0.5501	0.5024
Macao	0.9330	0.9171	0.9058	0.8903
Malaysia	0.3046	0.3092	0.6256	0.6351
Philippines	3.7656	3.8025	0.5316	0.5369
Sri Lanka	6.1945	5.9832	0.4794	0.4630
Vietnam	816.8254	823.3959	0.4006	0.4038

Table 10: Paasche-Laspeyres Spreads Based on the Official and Revised Basic Heading Data

Official Data

PLS	Bhutan	Fiji	Hong Kong	Indonesia	Macao	Malaysia	Philippines	Sri Lanka	Vietnam
Bhutan	1.0000	1.1743	1.6524	1.2227	1.7202	1.2779	1.1704	1.1835	1.0779
Fiji	1.1743	1.0000	1.1123	0.9974	0.9680	0.8525	0.9606	0.9647	1.0269
Hong Kong	1.6524	1.1123	1.0000	1.3605	0.9821	1.0874	1.4927	1.5025	1.8886
Indonesia	1.2227	0.9974	1.3605	1.0000	1.3880	1.1749	1.0842	1.1126	1.2365
Macao	1.7202	0.9680	0.9821	1.3880	1.0000	1.1108	1.4454	1.5362	1.9386
Malaysia	1.2779	0.8525	1.0874	1.1749	1.1108	1.0000	1.2159	1.2490	1.4509
Philippines	1.1704	0.9606	1.4927	1.0842	1.4454	1.2159	1.0000	1.1185	1.1956
Sri Lanka	1.1835	0.9647	1.5025	1.1126	1.5362	1.2490	1.1185	1.0000	1.1767
Vietnam	1.0779	1.0269	1.8886	1.2365	1.9386	1.4509	1.1956	1.1767	1.0000

Revised Data: CPRUD corrected for heteroscedasticty and semilogarithmic coefficient bias

PLS	Bhutan	Fiji	Hong Kong	Indonesia	Macao	Malaysia	Philippines	Sri Lanka	Vietnam
Bhutan	1.0000	1.2963	1.8988	1.5780	2.0014	1.5495	1.3295	1.4984	1.2795
Fiji	1.2963	1.0000	1.0952	1.0429	0.9514	0.9197	0.9540	1.0313	1.0276
Hong Kong	1.8988	1.0952	1.0000	1.4934	1.0048	1.1855	1.5011	1.5796	2.0391
Indonesia	1.5780	1.0429	1.4934	1.0000	1.5144	1.3338	1.1121	1.2754	1.3983
Macao	2.0014	0.9514	1.0048	1.5144	1.0000	1.1011	1.3852	1.5629	1.9324
Malaysia	1.5495	0.9197	1.1855	1.3338	1.1011	1.0000	1.2366	1.4033	1.4555
Philippines	1.3295	0.9540	1.5011	1.1121	1.3852	1.2366	1.0000	1.2795	1.2301
Sri Lanka	1.4984	1.0313	1.5796	1.2754	1.5629	1.4033	1.2795	1.0000	1.2937
Vietnam	1.2795	1.0276	2.0391	1.3983	1.9324	1.4555	1.2301	1.2937	1.0000