Basic Heading PPPs and the Lack of Matching Problem

Erwin Diewert[1] March, 2025.

1. Introduction

In these notes, we look at the problems associated with calculating interregional Purchasing Power Parity (PPP) indexes at the Basic Heading level when only price information is available. In particular, we deal with the case where not all products are priced in all countries. We focus on variants of the Country Product Dummy (CPD) method for linking countries or regions that was proposed by Summers (1973) and modified by Diewert (2004) to take into account that the World may be broken up into regions and the regions may want to impose regional fixity on the Basic Heading PPPs for the countries in their regions. Our discussion is loosely based on the methodology used by the International Comparison Program (ICP) to link the regional Basic Heading PPPs in 2017.[2] Our basic message is that the CPD method for calculating multilateral indexes is not reliable when there is a lack of matching of products across countries.

Section 2 develops the basic algebra for the single stage CPD multilateral method and section 3 does the same for the GEKS-Jevons method when there are missing prices. Readers familiar with multilateral methods can skip sections 2 and 3.

Section 4 develops the algebra of the corresponding two stage CPD and GEKS Jevons methods suggested by Diewert (2004) (2010) and Sergeev (2009) (2011) (2022). The two stage methods are needed because various regions use their own methods to form PPPs for countries withing a region. The job of the World Bank is to link the Regional PPPs into a set of global PPPs. We will focus on Sergeev's second stage method in section 5.

Section 5 provides some numerical examples of second stage CPD PPPs and Jevons multilateral PPPs using Basic Heading aggregate regional average price data that were collected for the ICP in 2017.

Section 6 concludes with some recommendations and notes some additional problems with the ICP that should be discussed.

2. The Prices Only Algebra for the Country Product Dummy Multilateral PPPs

It is useful to present the algebra for the CPD method when there are missing prices.

Suppose that there are K countries in a regional international comparison of prices with N products in scope. If product n is priced in country k for the time period under consideration, denote its average price by p_{kn} for $k = 1,\ldots,K$ and $n \in S(k)$ where $S(k)$ denotes the set of products that are priced in country k. The price is expressed in units of country k's currency.[3] The basic assumption made in the CPD model is that the observed country prices satisfy the following equations (approximately):

$$(1)\ p_{kn} \approx \pi_k \alpha_n ; \qquad k = 1,\ldots,K;\ n \in S(k);$$

---

[1] The author thanks Yuri Dikhanov for many valuable suggestions. However, he does not necessarily agree with some parts of the paper and is not responsible for any mistakes made by the author.

[2] The ICP methodology makes use of importance weighting which we ignore in these notes for simplicity. Thus the numerical results presented in these notes that make use of ICP country BH data for products do not reflect actual ICP computations.

[3] Each product is priced in a common unit of measurement across countries.

where α_n is a quality adjustment parameter for product n and π_k is the overall level of prices (for the N product prices in scope) in country k relative to the level of prices in other countries. Thus the basic assumption is that product prices vary proportionally across countries for the group of products in scope. It is expected that the π_k are approximately proportional to country k's exchange rate for k = 1,…,K. The parameter π_k is country k's Purchasing Power Parity.

Take logarithms of both sides of equations (1) and add error terms to obtain the following linear regression model:

(2) $\ln p_{kn} = \rho_k + \beta_n + \varepsilon_{kn}$;

$k = 1,\ldots,K;\ n\in S(k)$;

where ε_{kn} is an error term and ρ_k and β_n are the logarithms of π_k and α_n; i.e.,

(3) $\rho_k \equiv \ln\pi_k$ for $k = 1,\ldots,K$ and $\beta_n \equiv \ln\alpha_n$ for $n = 1,\ldots,N$.

Estimates for the parameters ρ_k and β_n in equations (2) can be found by solving the following least squares minimization problem:

(4) $\min_{\rho,\beta} \Sigma_{k=1}^{K} \Sigma_{n\in S(k)} [\ln p_{kn} - \rho_k - \beta_n]^2 = \min_{\rho,\beta} \Sigma_{n=1}^{n} \Sigma_{k\in S^*(n)} [\ln p_{kn} - \rho_k - \beta_n]^2$

where $\rho \equiv [\rho_1,\ldots,\rho_K]$ and $\beta \equiv [\beta_1,\ldots,\beta_N]$ and $S^*(n)$ is the set of countries k that have priced product n for n = 1,…,N. Note that there are two equivalent ways of writing the least squares minimization problem. Solutions ρ and β to problem (4) will satisfy the following first order conditions for solving the minimization problem:

(5) $\Sigma_{n\in S(k)} \ln p_{kn} = \Sigma_{n\in S(k)} [\rho_k + \beta_n]$; $k = 1,\ldots,K$;
(6) $\Sigma_{k\in S^*(n)} \ln p_{kn} = \Sigma_{k\in S^*(n)} [\rho_k + \beta_n]$; $n = 1,\ldots,N$.

Let N(k) equal the number of products that are priced in country k for k = 1,…,K and let $N^*(n)$ equal the number of countries that price product n for n = 1,…,N. Using these definitions plus definitions (3) (to switch from the ρ_k and β_n to the π_k and α_n), equations (5) and (6) can be rewritten as follows:[4]

(7) $\pi_k = \Pi_{n\in S(k)} [p_{kn}/\alpha_n]^{1/N(k)}$; $k = 1,\ldots,K$;
(8) $\alpha_n = \Pi_{k\in S^*(n)} [p_{kn}/\pi_k]^{1/N^*(n)}$; $n = 1,\ldots,N$.

A solution $\pi \equiv [\pi_1,\ldots,\pi_K]$ and $\alpha \equiv [\alpha_1,\ldots,\alpha_N]$ to equations (7) and (8) will not be unique since $\lambda\pi$ and $\lambda^{-1}\alpha$ will also be a solution for any positive scalar λ. Thus we are allowed to make one normalization on the π_k and α_n in order to obtain a unique solution. If we choose the normalization $\pi_1 = 1$, then we are choosing country 1 as the world numeraire country. From equations (8), it can be seen that in addition to the interpretation of the α_n as quality adjustment parameters, α_n can be interpreted as a PPP adjusted regional average price for product n.

A fundamental problem with the Purchasing Power Parities π_k defined by a solution to the least squares minimization problem (4) when there are missing prices is the fact that the ratio of the PPP for country k to

---

[4] The sets S(k) and $S^*(n)$ are assumed to be nonempty for k = 1,…,K and n = 1,…,N. For more details on the algebra in this section, see Rao (1995) (2005) and Diewert (2004) (2023).

the PPP for country m may not depend on any matched product prices between the two countries; i.e., we have:

(9) $\pi_k/\pi_m = \Pi_{n\in S(k)} [p_{kn}/\alpha_n]^{1/N(k)}/\Pi_{n\in S(m)} [p_{mn}/\alpha_n]^{1/N(m)}$.

But the two countries may not have a single product in common so the product prices in the numerator of the right hand side of (9) may be totally different from the product prices in the denominator of the right hand side of (9). Thus the CPD multilateral method with missing prices will in general have a *lack of matching problem*.[5]

There is another problem with the CPD approach to multilateral index number theory that arises if we take the economic approach to index number theory: the CPD indexes implicitly assume that purchasers of the N products *have linear preferences over the products in scope*.[6] Equations (1) which form the foundation for the CPD model are consistent with purchasers in a country maximizing the linear utility function, $f(q_1,\ldots,q_N) \equiv \Sigma_{n=1}^{N} \alpha_n q_n$ (subject to budget constraints) where the marginal utility parameters α_n appear in equations (1) and the q_n are total market purchases of product n for the country under consideration.[7] The underlying economic model which justifies equations (1) is a model where consumers regard the N products as *perfect substitutes* (after quality adjustment). The assumption that a group of products are (approximately) perfect substitutes is satisfactory if the products in scope are closely related. For example, in the current ICP list of product groups, one group or Basic Heading category in the International Comparison Project (ICP) is “Rice” and another is “Fresh Milk”. The assumption that the detailed products in these groups are (approximately) perfect substitutes is probably a satisfactory one. But consider the Basic Heading (BH) “Jewelry, Clocks and Watches” [8] or the BH “Cultural Services”.[9] “Pharmaceuticals” is another interesting product category.[10] There are thousands of detailed products in each of these groups and many of them are not close to being perfect substitutes. Thus from the perspective of the economic approach to index number theory, the CPD approach to multilateral index number theory is not entirely satisfactory if the product categories are broad.

A final potential problem with the CPD approach to multilateral index number theory concerns the number of degrees of freedom relative to the number of parameters in the model. As will be seen in section 5, it can happen that for some product categories, there may be very few degrees of freedom per parameter due to missing prices in many countries.

---

[5] To further illustrate the lack of matching problem with the CPD method, consider an example with K = 2 and N = 2. Product 1 is only available in country 1 and product 2 is only available in country 2. Solutions to the CPD equations (7) and (8) plus the equation $\pi_1 = 1$ in this case boil down to $\alpha_1 = p_{11}$; $\alpha_2 > 0$ is arbitrary, $\pi_1 = 1$ and $\pi_2 = p_{22}/\alpha_2$. The price index between the two countries is $\pi_2/\pi_1 = p_{22}/\alpha_2$. Since α_2 can be any positive number, we see that the price index between the two countries is not well defined in this case.

[6] The CPD indexes are also consistent with purchasers having Cobb-Douglas preferences over all products in scope. But Cobb-Douglas preferences are not consistent with the existence of mission products or zero expenditure shares on products in any country.

[7] See section 5 in Diewert (2022) for the details of this economic model and its connection to hedonic regressions. The assumption that purchasers have the same linear preferences is not so problematic if the CPD model fits the data well *and* there are a sufficient number of degrees of freedom per parameter.

[8] Here are the 5 detailed products that the ICP prices across countries for this BH group of products: (1) Wrist-watch, children's, SWATCH Flik Flak; (2) Wrist-watch, men's, CITIZEN Eco-Drive BM6060; (3) Analog travel alarm, quartz; (4) Wedding ring, 14 Karat gold and (5) Wall clock, SEIKO.

[9] Here are the 4 detailed products that the ICP prices across countries for this BH group of products: (1) Digital photo print; (2) Cinema ticket, weekend; (3) Passport-type photos, by photographer and (4) Rental of DVD movie.

[10] It is obvious that different drugs are far from being perfect substitutes for each other.

For the 2017 ICP, there were 5 separate regions in the ICP: (i) Africa with 50 countries; (ii) Asia with 22 countries; (iii) the OECD with 49 countries; (iv) Latin America with 13 countries and (v) Western Asia with 12 countries. The five regions prepared regional average prices for 86 Basic Heading (BH) categories. The number of specific products in the ICP's core product list was 631.The number of specific products in each BH grouping ranged from 1 to 57 but the median number of products in each BH was 5. The 5 regions (in theory) supplied regional average prices to the World Bank for each of the 631 specific products. In practice, the regions supplied regional average prices for only about 2/3 of the core list of 631 products. Here are the number of core list average product prices supplied to the ICP by region: 415, 461, 407, 415, 440. It is possible to apply the CPD method using regional average prices for each BH group of products in order to link the price levels across the regions for each BH group of products. This is the second stage of the two stage procedure used by the World Bank to link the regions. We will discuss this second stage linking problem in more detail below where we will see that the problem of missing prices causes problems. When one looks at equation (9) in the context where many prices are missing, it can be seen that the CPD PPPs are basically meaningless when the Basic Heading category contains products that are very different. The CPD method works well when there is a great deal of matching of prices across countries.

If each country prices all N products, then equations (7) and (8) simplify to the following equations:

$$(10)\ \pi_k = \Pi_{n=1}^{N} [p_{kn}/\alpha_n]^{1/N}\ ; \qquad k = 1,\ldots,K;$$
$$(11)\ \alpha_n = \Pi_{k=1}^{K} [p_{kn}/\pi_k]^{1/K}\ ; \qquad n = 1,\ldots,N.$$

Thus in this case, the PPP ratio of country k relative to county m, becomes $\pi_k/\pi_m = \Pi_{n=1}^{N} [p_{kn}/p_{mn}]^{1/N}$ which is the bilateral Jevons index between countries k relative to m.[11] This special case is an important one; *there is no lack of matching problem in this case*. If each country prices all N products, then the system of CPD country Purchasing Power Parities PPP_k will be proportional to the simple geometric means of the N product prices for each country; i.e., we can set the PPP_k equal to the geometric mean of each country's prices for the N products:

$$(12)\ PPP_k = \Pi_{n=1}^{N} [p_{kn}]^{1/N}\ ; \qquad k = 1,\ldots,K.$$

In the general case when there are missing prices, there is another multilateral method that we want to consider and that is the Gini or GEKS-Jevons multilateral method. We write out the algebra for this method in the following section.

3. The Maximum Overlap "Star" Jevons indexes and the GEKS Jevons Index

Define S(k,m) as the set of products that are priced in countries k and m[12] and define N(k,m) as the number of common products that are priced in countries k and m for k = 1,…,K and m = 1,…,K. The *maximum overlap Jevons index* which compares the common product prices in country m to the corresponding product prices in country k, $P_J(m/k)$, is defined as the geometric mean of the common product price ratios:

$$(13)\ P_J(m/k) \equiv [\Pi_{n\in S(k.m)}\, (p_{mn}/p_{kn})]^{1/N(k,m)}\ ; \qquad m = 1,\ldots,K;\ k = 1,\ldots,K.$$

The K vectors of K *maximum overlap Jevons star indexes*, $P_J(k)$, are defined as follows:

[11] This result is a special case of a more general result obtained by Triplett and McDonald (1977; 150). The Jevons index has the "best" axiomatic properties for prices only index number formulae that do not have missing prices; see Diewert (1995). In particular, the Jevons index with no missing products satisfies the circularity test.

[12] We assume these K^2 sets are not empty. For certain narrowly defined categories, it could happen that countries k and m have no products in common. In this case, we cannot define a meaningful price index relating the prices of country k to the prices of country m.

(14) $P_J(k) \equiv [P_J(1/k), P_J(2/k), \ldots, P_J(K/k)]$; $k = 1,\ldots,K.$

Advantages of a Jevons star index over a CPD index are its simplicity, ease of computation and relevance. A Jevons bilateral star index simply takes an average of price ratios for identical products across countries rather than on estimating preferences of purchasers across countries. The focus is on obtaining product matches across countries; the greater the number of matches, the more reliable the index will be. A disadvantage of the Jevons star methodology is that in the case of missing observations, there are K possible choices of a base country and the different Jevons star PPPs will in general generate different PPPs.

Following the example of Gini (1931) and others, the vector of *GEKS-Jevons indexes*, P_{GEKS-J}, is defined (up to a proportional factor) by taking the geometric mean of the Jevons star indexes. The resulting indexes turn out to be proportional to the following vector:

(15) $P_{GEKS-J} \equiv [\Pi_{k=1}^{K} P_J(1/k)^{1/K} , \Pi_{k=1}^{K} P_J(2/k)^{1/K},\ldots, \Pi_{k=1}^{K} P_J(K/k)^{1/K}].$

As noted above, if there are missing prices, then the resulting Jevons star indexes can differ substantially from each other. The GEKS Jevons index simply takes an average of the K star indexes. In the case where the individual star indexes are far from being proportional, then the reliability of the GEKS Jevons index is questionable.[13] However, *if there are no missing prices* across the N products and K countries, then each K dimensional vector of Jevons star indexes, $P_J(1)$, $P_J(2)$, …., $P_J(K)$ and the GEKS-Jevons index P_{GEKS-J} are all proportional to each other and to the vector of Jevons indexes P_J whose components are defined by (12). Thus if there are no missing prices, then all of the Jevon star indexes, the GEKS-Jevons index and the CPD vector of PPPs are all proportional to P_J defined as follows:

(16) $P_J \equiv [\Pi_{n=1}^{N} [p_{1n}]^{1/N}, \Pi_{n=1}^{N} [p_{2n}]^{1/N},\ldots, \Pi_{n=1}^{N} [p_{Kn}]^{1/N}].$

Component 1 of P_J is the simple geometric mean of the N product prices that are priced in country 1, component 2 of P_J is the simple geometric mean of the same N products that are priced in country 2 and so on. It is easy to explain to the public the intercountry Jevons PPPs when all N products are priced by every country.

The conclusion that can be drawn from this section and the previous one is this: *it is a very good idea to design the collection of international prices across countries so that there are no missing prices*. In the *no missing prices case*, *all K+2 indexes defined in these two sections are proportional to each other and hence equal to each other* if we make the same country the numeraire country across the various index number concepts.

4. The ICP Methodology for Linking the Regions

The 2017 ICP methodology for linking the regions is not a straightforward application of the CPD methodology described in the previous sections. A very incomplete outline of the ICP methodology[14] is presented below.

[13] An extreme example of unreliability of the GEKS Jevons index arises if there are so many missing prices that one or more maximum overlap bilateral Jevons indexes cannot be calculated because there are no common products between the two countries. This extreme case can actually happen as we shall see later.

[14] The ICP methodology made use of importance weights in the prices that are collected to link the regions. The importance weights took on the value 1 (not so important) or 3 (the product is important in the home country). These weights are unlikely to be determined with precision. Thus the exposition below does not make use of importance

The world was divided up into 5 regions:

- Region 1 was Africa with 50 countries;
- Region 2 was Asia with 22 countries;
- Region 3 was the OECD with 49 countries;
- Region 4 was Latin America with 13 countries and
- Region 5 was West Asia with 12 member countries.

There were also several singleton countries that did not belong to any region and a few countries that belonged to two regions. Roughly speaking, there were 86 Basic Heading consumption categories of product that were used to compare prices across the 5 regions.[15] Within each of these Basic Heading categories, the ICP, in cooperation with the regions, assembled a core list of 631 specific products were priced (in the same units of measurement) across the regions so that the prices of the regions could be linked to each other. The number of specific products within each BH category varied between 1[16] and 57.[17] There were two categories with only 2 specific products.[18] The median number of specific products in each BH category was 5. Once the country prices were collected, they were divided by the regional PPPs for the respective Basic Heading categories. The regionally deflated prices were then used to link the regional PPPs into World PPPs using a variant of the Country Product Dummy model.

The missing prices problem is a big one. If each country had priced all 631 products, the total number of prices collected for comparison purposes across all 146 countries would be 92,126. The number of products that were actually priced was 40,949.[19] Thus the overall percentage of prices collected over possible prices collected was 44.4%. Table 1 gives a breakdown of the percentage of Basic Heading prices collected by region.

Table 1: Prices Collected versus Maximum Possible Prices Collected by Region in 2017

	All Regions	Africa	Asia	OECD	Latin America	Western Asia
Maximum # of Prices Possible	**92126**	**31550**	**13882**	**30919**	**8203**	**7572**
Actual # of Prices Collected	**40949**	**17948**	**7118**	**7856**	**3239**	**4788**
Percentage of Prices Collected	**44.4**	**56.9**	**51.3**	**25.4**	**39.5**	**63.2**

We now describe possible methods for linking the regions. As indicated above, in the International Comparison Program (ICP), the regions construct their own set of PPPs for each Basic Heading Category. For various reasons, the regions do not want the method which links the PPPs of one region to the PPPs of another region to be affected by the ICP's method for linking the regions; i.e., regional BH PPPs need to

weights and thus does not fully describe the ICP methodology. However, the lack of matching problem does not go away even if importance weights are used in the weighted CPD method.

[15] However, for 22 Basic Heading categories, no prices were collected; exchange rates or other exogenous indexes were used to fill in the PPPs for these categories.

[16] The BH category that has only one specific product is the category "Household Services". There are two specific products listed in this class of products: (i) Laundry and (ii) Laundry; Self Service. However, not a single country in the world priced the second product so there is only one effective product in this (very broad) BH category.

[17] The BH category that has 57 specific products is "Pharmaceuticals".

[18] The BH category "Eggs and Egg Based Products" has two specific products: (i) Chicken eggs, caged hen, large size and (ii) Chicken eggs, caged hen, medium size. The BH category "Sugar" has two specific products: (i) White sugar and (ii) Brown sugar.

[19] Again, this description does not do justice to the ICP methodology.

be respected by the interregional linking method. Thus we will consider two methods for linking the regions that respect regional Basic Heading PPPs.

We introduce some notation for the regional Basic Heading PPPs. Suppose that there are 5 regions that participate in an international comparison of consumer prices. Each region r has C(r) countries in it for r = 1,…,5. Also, as noted above, for the ICP, there are 631 Basic Heading categories for products and we focus on one of them and develop the algebra below for this chosen BH category. Within this given BH category, we assume that there are N items on the world core list of products in the chosen category that every country in the comparison will endeavor to price during the comparison year. Not every country will be able to price every item on this core list of products. Thus for each region r and each country c within that region, there will be a subset of the core product list, S(r,c), that country c in region r will be able to price in units of its domestic currency. For such products n where $n \in S(r,c)$, this price will be denoted as p_{rcn}, where r = 1,...,5, c = 1,...,C(r) and $n \in S(r,c)$ where there are C(r) countries in region r. Within each region, we assume that country Basic Heading PPPs for the basic heading class of products under consideration has been determined by the regional coordinator for the relevant region. The Basic Heading PPP for country c in region r is denoted by γ_{rc} for r = 1,...,5 and c = 1,...,C(r).[20] These regional parities are used to convert the item prices into *common regional currency prices*, p_{rcn}^*, defined as follows:

$$(17)\ p_{rcn}^* \equiv p_{rcn}/\gamma_{rc}\ ; \qquad r = 1,\ldots,5;\ c = 1,\ldots,C(r);\ n \in S(r,c).$$

The following assumption is made: the normalized regional prices defined by (17) satisfy the following CPD equations (approximately):

$$(18)\ p_{rcn}^* \approx \pi_r\alpha_n\ ; \qquad r = 1,\ldots,5;\ c = 1,\ldots,C(r);\ n \in S(r,c).$$

The π_r are the desired interregional PPPs and the α_n are the usual quality adjustment parameters.[21] Define the parameters ρ_r and β_n as the logarithms of the π_r and α_n:

$$(19)\ \rho_r \equiv \log\pi_r\ ;\ r = 1,\ldots,5;\ \ \beta_n \equiv \log\alpha_n\ ;\ n = 1,\ldots,N.$$

Now take logarithms of both sides of equations (18), add error terms to the resulting equations and we obtain a linear regression model. Thus solutions ρ_r and β_n to the resulting linear regression problem can be found by solving the following least squares minimization problem:

$$(20)\ \min_{\rho,\beta}\ \{\Sigma_{r=1}^{5}\ \Sigma_{c=1}^{C(r)}\ \Sigma_{n\in S(r,c)}\ [\ln p_{rcn}^* - \rho_r - \beta_n]^2\} = \min_{\rho,\beta}\ \{\Sigma_{n=1}^{N}\ \Sigma_{r=1}^{5}\ \Sigma_{c\in S^*(r,n)}\ [\ln p_{rcn}^* - \rho_r - \beta_n]^2\}$$

where $S^*(r,n)$ is the set of all countries in region r that price product n. The first order necessary conditions for $\rho \equiv [\rho_1,\ldots,\rho_5]$ and $\beta \equiv [\beta_1,\ldots,\beta_N]$ to solve (20) can be written as follows:

$$(21)\ \Sigma_{c=1}^{C(r)}\ \Sigma_{n\in S(r,c)}\ \ln p_{rcn}^* = \Sigma_{c=1}^{C(r)}\ \Sigma_{n\in S(r,c)}\ [\rho_r + \beta_n]\ ; \qquad r = 1,\ldots,5;$$
$$(22)\ \Sigma_{r=1}^{5}\ \Sigma_{c\in S^*(r,n)}\ \ln p_{rcn}^* = \Sigma_{r=1}^{5}\ \Sigma_{c\in S^*(r,n)}\ [\rho_r + \beta_n]\ ; \qquad n = 1,\ldots,N.$$

In order to proceed further, we need to make a few more definitions:

(23) $N(r,c) \equiv$ the number of products priced in country c of region r; r = 1,…,5; c = 1,…,C(r);
(24) $N(r) \equiv \Sigma_{c=1}^{C(r)} N(r,c)$ is the total number of products priced in region r; r = 1,…,5;

[20] In each region r, one country will be chosen as the numeraire country and if that country is country 1, then γ_{r1} is set equal to 1.
[21] This methodology was originally suggested by Diewert (2004) (2010).

(25) $N^*(r,n) \equiv$ the number of countries in region r that priced product n; r = 1,…,5; n = 1,…,N;
(26) $N^*(n) \equiv \Sigma_{r=1}^{5} N^*(r,n)$ is the total number of countries that priced product n ; n = 1,…,N.

Using definitions (19) and (23)-(26), equations (21) and (22) can be transformed into the following equations that a solution to (20) must satisfy:

(27) $\pi_r = \{\Pi_{c=1}^{C(r)} \Pi_{n\in S(r,c)}[p_{rcn}^*/\alpha_n]\}^{1/N(r)}$; r = 1,…,5;
(28) $\alpha_n = \{\Pi_{r=1}^{5} \Pi_{c\in S^*(r,n)}[p_{rcn}^*/\pi_r]\}^{1/N^*(n)}$; n = 1,…,N.

Thus the interregional PPP for region r, π_r, is the geometric mean of all of the quality adjusted prices for products priced in region r and the nth world wide quality adjustment factor for product n is the geometric mean of all regional deflated product n prices p_{rcn}^* over all regions deflated by the appropriate regional PPP for each region, π_r.[22]

Recall that the prices p_{rcn}^* are normalized prices p_{rcn}/γ_{rc}; i.e., they are the country prices divided by the regional PPP for the BH category in scope. Thus if we replace the p_{rcn}^* by the p_{rcn}/γ_{rc}, equations (27) and (28) become the following equations;

(29) $\pi_r = \{\Pi_{c=1}^{C(r)} \Pi_{n\in S(r,c)}[p_{rcn}/\gamma_{rc}\alpha_n]\}^{1/N(r)}$; r = 1,…,5;
(30) $\alpha_n = \{\Pi_{r=1}^{5} \Pi_{c\in S^*(r,n)}[p_{rcn}/\gamma_{rc}\pi_r]\}^{1/N^*(n)}$; n = 1,…,N.

The π_r for r = 1,…,5 are the regional price levels for the group of products in the Basic Heading category under consideration. The interregional PPP for region r relative to region t is the ratio π_r/π_t which is equal to the following expression using equations (29):

(31) $\pi_r/\pi_t = \{\Pi_{c=1}^{C(r)} \Pi_{n\in S(r,c)}[p_{rcn}/\gamma_{rc}\alpha_n]\}^{1/N(r)}/\{\Pi_{c=1}^{C(t)} \Pi_{n\in S(t,c)}[p_{tcn}/\gamma_{tc}\alpha_n]\}^{1/N(t)}$.

A variation of this method was first used in the 2005 ICP comparison using a subset of all countries (18 Ring countries) to compare prices across regions. For 2011 and 2017, the concept was modified in two important ways: (1) all countries could price the global common list of products, i.e., in principle, all countries could participate in linking the regions instead of just using 18 Ring countries; and (2) importance weights {3,1} were introduced. We will not go into more detail here, as it is not important for our exposition.

There are three major problems with the interregional PPP ratios defined by (31) in the case where there are missing prices:

- As was the case in the straightforward CPD method, *with missing observations*, the prices for product n in the numerator and denominator of (31) will not be perfectly matched and in extreme cases, there could be no matched prices on the right hand side of (31). Since index number theory basically rests on aggregating price ratios for matched products, it is difficult to regard the PPPs defined by (31) as being meaningful in the case where there are a low number of product matches across the regions.
- Looking at (30), it is apparent that regions with a large number of countries will tend to have more prices collected for product n and hence, the large regions will tend to determine the relative magnitudes of the α_n. But the α_n play a large role in the determination of the interregional price levels, π_r as is indicated by equations (29). Thus this methodological approach to linking the regions is not "democratic" since regions with a larger number of countries will tend to play a larger role

[22] As usual, the solution is not unique. In order to obtain a unique solution, we need to add a normalization like π_1 = 1.

in the determination of the interregional PPPs.[23] This problem persists even if each country in each region prices all N products; i.e., this problem is independent of the lack of matching problem.

- The regional PPPs, the γ_{rc}, are not necessarily the "right" regional PPPs to deflate the N particular products chosen by the World Bank and the regional coordinators to be representative for all the hundreds of products that are contained in each Basic Heading category. However, finding the "true" deflators for all countries within a region is a more or less impossible task. Even if we could find the "true" within region PPPs, it is not clear how this information would lead to better interregional PPPs.

In order to illustrate potential lack of matching problems when making international comparisons, we will look at some examples drawn from the ICP's 2017 data. There were 631 products on the Global Core List in 2017; however some items were not priced in any region and therefore should be excluded from the calculations. On top of that, some additional items were effectively excluded from computing inter-regional PPPs as they were only priced in one region. Altogether, there were 50 items like that, so the actual number of items used in calculating the inter-regional PPPs was 581 and the overall fill rate was 72.2%, ranging from 67.8% (OECD) to 77.1% (Asia):

Table 2: Products Priced by the Regions in 2017

	All Regions	Africa	Asia	OECD	Latin America	Western Asia
Maximum # of Prices Possible	**2905**	**581**	**581**	**581**	**581**	**581**
Actual # of Prices Collected	**2097**	**409**	**448**	**394**	**413**	**433**
Percentage of Prices Collected	**72.2**	**70.4**	**77.1**	**67.8**	**71.1**	**74.5**

While the average fill rates are relatively high, they are not uniform, and some Basic Headings have much lower rates. More importantly, there could be a significant variance in fill rates between regions in the same Basic Heading, where the fill rates can range from 25% to 100%.

As noted above, the regional coordinators provided the ICP with region wide average prices for most of the products listed in the ICP's core product list for the particular Basic Heading category under consideration. Thus let S(r) be the set of products n in the BH category under consideration that are priced by region r and let N(r) be the number of average product prices that are provided to the ICP from region r for r = 1,…,5. Let $S^*(n)$ be the set of regions that price product n and let $N^*(n)$ equal to the number of regions that price product n for n = 1,…,N. The region r supplies the ICP with the regional average price for product n in the Basic Heading product category under consideration, P_{rn}, for r = 1,…,5 and $n \in S(r)$. These regional average prices for product n are expressed in units of the currency of the numeraire country in each region.[24] Below is a Table listing the numeraire countries for the five regions for the 2017 round of the ICP.

Table 3: Regional Numeraire Countries

Region	Name	Country Numbers	Numeraire Country
1	**Africa**	**1-50**	**South Africa (#48)**
2	**Asia**	**51-72**	**Hong Kong (#56)**
3	**OECD**	**73-121**	**USA (#121)**
4	**Latin America**	**122-134**	**Brazil (#124)**
5	**Western Asia**	**135-146**	**Oman (#142)**

[23] This problem with the methodology was first noticed by Sergeev (2009) (2011) (2022).

[24] It is not a trivial job for the regions to prepare these regional average prices. There are many problems in getting annual average prices for a product for a single country and the aggregation over countries problem is difficult.

We turn now to some methodological issues raised by the fact that the above method for linking the regions gives more weight to regions which have more countries.

Sergeev (2009) proposed an alternative more "democratic" method for linking the regions and we will now describe his method for linking the regions. Basically, his method applied to a Basic Heading category is to apply the CPD method using the *regional average product prices* P_{rn} instead of the (deflated) country prices for product n, p_{rcn}/γ_{rc}. [25]

The basic model assumption for Sergeev's Democratic Model is the following counterpart to the CPD model described by (1) in section 1:

$$(32)\ P_{rn} \approx \pi_r\alpha_n\ ; \qquad r = 1,\ldots,5;\ n\in S(r)$$

where S(r) is the set of products n that were available in region r.

Making the same assumptions and definitions as were made in section 1, we find that the $\pi \equiv [\pi_1,\pi_2,\pi_3,\pi_4,\pi_5]$ and $\alpha \equiv [\alpha_1,\ldots,\alpha_N]$ solution to the least squares minimization problem for the model defined by that is the counterpart to problem (4) must satisfy the following equations:

$$(33)\ \pi_r = \Pi_{n\in S(r)}\,[P_{rn}/\alpha_n]^{1/N(r)}\ ; \qquad r = 1,\ldots,5;$$
$$(34)\ \alpha_n = \Pi_{r\in S^*(n)}\,[P_{rn}/\pi_r]^{1/N^*(n)}\ ; \qquad n = 1,\ldots,N.$$

The ratio of the price level for region r π_r relative to the price level for region t π_t is:

$$(35)\ \pi_r/\pi_t = \Pi_{n\in S(r)}\,[P_{rn}/\alpha_n]^{1/N(r)}/\ \Pi_{n\in S(t)}\,[P_{tn}/\alpha_n]^{1/N(t)}.$$

If regions r and t price the same products, then S(r) = S(t) and N(r) = N(t) and (35) becomes the following maximum overlap bilateral Jevons index between regions r and t:

$$(36)\ \pi_r/\pi_t = \Pi_{n\in S(r)}\,[P_{rn}/P_{tn}]^{1/N(r)}.$$

Sergeev's method is indeed much more democratic than Diewert's method but not all regions price all products in a Basic Heading category and this allows the regions which price more products to have a bigger role in the interregional comparison of prices. However, if all regions provided a complete set of regional average prices (the P_{rn}) to their regional coordinators, then Sergeev's average price CPD method would be fully democratic; i.e., each region would have exactly the same influence on the interregional PPPs. In any case, Sergeev's method is indeed much more democratic than Diewert's method.

However, Sergeev's method can still suffer from a missing price problem which can be very acute for some Basic Heading categories as we shall see in the next section.

5. Some Examples of Interregional Linking Using Sergeev's Method

---

[25] In principle, if each country c in region r provided the regional coordinator with the product n price p_{rcn}, then P_{rn} should equal the geometric mean of these (deflated by within region country PPPs) prices; i.e., $P_{rn} = [\Pi_{c=1}^{C(r)}(p_{rcn}/\gamma_{rc})]^{1/C(r)}$.

We will present 5 examples of interregional linking using the ICP 2017 BH data for the Regional average prices for a few of the ICP Basic Heading product categories.[26] We calculate the Sergeev's CPD index as well as the Jevons Star indexes and the GEKS-Jevons index for each of our 5 regions for the particular Basic Heading category under consideration.

However, it should be kept in mind that the following examples do not represent actual calculations, as they follow neither the 2005 nor 2011-17 methodologies and are presented for illustrative purposes only. However, they are perfectly adequate for the purposes of our exposition that attempts to show that it is important to have matched prices across regions. This section is relevant both for calculating interregional PPPs and individual country PPPs within a region.

Example 1: Basic Heading Category is Rice

There are 11 products in the Rice BH category of products. Here is a listing of the individual products along with their average prices by region:

Table 4: Regional Geometric Average Prices for the BH Category Rice

Product title	Region 1	Region 2	Region 3	Region 4	Region 5
Long-grain rice, parboiled, WKB	15.48997	11.90368	2.50671	4.69858	0.88093
Long-grain rice, not parboiled, WKB	16.37012	12.49387	1.14681	4.17688	0.94589
Long grain rice, family pack, WKB	17.16246	0	0	0	0
Jasmine rice, WKB	0	0	0	0	0
Basmati rice, WKB	33.32396	22.13316	4.01875	20.28345	1.13046
Broken rice, 25%, BNR	11.29647	7.02469	0	2.91915	0
Medium-grain rice, BNR	12.44030	8.84751	0	3.65276	0
Brown rice, family pack, BL	0	13.38436	0	0	0
Short-grain rice, BNR	11.78141	8.19426	0	0	0
Sticky rice, WKB	0	13.23967	4.95208	0	0
Long-grain rice, UNCLE BEN'S ORIGINAL	0	0	0.00000	0	0

Products 4 and 11 were not priced in any country (or at least, average regional prices for these products were not calculated). Thus they are dropped from the list of products that are used in the CPD regression model. Products 3 and 8 were only priced in one country so these products can only have a very limited contribution to the calculation of interregional price indexes.[27] The maximum overlap Jevons indexes will not use these isolated prices.

The interregional CPD, GEKS-Jevons (GEKS-J) and Jevons maximum overlap star indexes, J1-J5, are listed in Table 5 below. These indexes have been normalized so that the OECD level is equal to 1 for all 7 indexes.

Table 5: Alternative Interregional PPPs for the BH Rice

Region	P_{CPD}	$P_{GEKS\text{-}J}$	P_{J1}	P_{J2}	P_{J3}	P_{J4}	P_{J5}
1	8.229	8.872	9.010	7.482	9.010	10.041	9.010
2	5.445	6.052	6.326	5.254	5.254	7.064	6.580
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	2.680	3.001	2.920	2.420	3.254	3.254	3.254
5	0.385	0.415	0.434	0.346	0.434	0.434	0.434

[26] As noted above, the ICP methodology for linking the regions is more complicated than the methods that we explain below. However, the examples below use ICP data for the 2017 ICP. The prices listed in the examples are regional average prices so no individual country prices appear in the examples.
[27] If products 3 and 8 are dropped from the CPD regression, the resulting PPPs are the same as the PPPs generated by the full set of 9 products.

It can be seen that the 5 Jevons star indexes differ considerably and the CPD and GEKS-Jevons indexes also exhibit a considerable amount of variation. This variation is caused by the missing prices: if each region priced all 7 overlapping products, all 7 of the indexes listed in Table 5 would be identical.

Example 2: Basic Heading Category is Other Cereals, Flour and Other Cereal Products

There are 5 products in the Cereals BH category of products. Here is a listing of the individual products along with their average prices by region:

Table 6: Regional Average Prices for the BH Other Cereals

Product title	Region 1	Region 2	Region 3	Region 4	Region 5
Cornflakes, KELLOGG'S	99.31355	117.92527	6.36323	37.28978	2.46250
Wheat flour, not self-rising, BL	11.32623	11.89930	0.53268	4.49298	0.34202
Wheat semolina (suji), WKB	0	15.47009	1.65450	0	0.56756
Oats, rolled, WKB	0	0	2.55555	13.72639	0
Corn (maize) flour, white, WKB	12.10363	22.88166	0	0	0.53977

Only 18 of the 25 possible regional average prices are positive. This means that we will have only 18 degrees of freedom for the CPD model, which has 5 + 5 – 1 = 9 parameters. Thus we have only 2 degrees of freedom for each parameter.

The interregional CPD, GEKS-Jevons (GEKS-J) and Jevons maximum overlap star indexes, J1-J5, are listed in Table 7 below. These indexes have been normalized so that the OECD level is equal to 1 for all 7 indexes.

Table 7: Alternative Interregional PPPs for the BH Other Cereals

Region	P_{CPD}	P_{GEKS-J}	P_{J1}	P_{J2}	P_{J3}	P_{J4}	P_{J5}
1	13.66883	15.48717	18.21698	11.79602	18.21698	16.65341	13.66686
2	16.49907	17.87091	24.24814	15.70138	15.70138	18.60031	16.39303
3	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
4	5.86518	6.28162	7.03059	5.42545	6.42715	6.42715	6.20715
5	0.44294	0.46536	0.58661	0.42152	0.44009	0.45569	0.44009

It can be seen that the 5 Jevons star indexes vary enormously and the CPD and GEKS-Jevons indexes also exhibit a considerable amount of variation. This variation is caused by the missing prices: again, if each region priced all 5 products, all 7 of the indexes listed in Table 7 would be identical.

Example 3: Basic Heading Category is Eggs and Egg-Based Products

There are 2 products in the Eggs BH category of products. Here is a listing of the individual products along with their average prices by region:

Table 8: Regional Average Prices for the BH Category Eggs

Product title	Region 1	Region 2	Region 3	Region 4	Region 5
Chicken eggs, caged hen, large size	13.44271	11.95967	1.133313	3.419534	0.396964
Chicken eggs, caged hen, medium size	0	10.30738	1.001608	3.028731	0.378172

Only 1 price out of 10 possible regional average prices is missing. There are 9 degrees of freedom and 6 CPD parameters to estimate so the number of degrees of freedom per parameter is 1.5.

The interregional CPD, GEKS-Jevons (GEKS-J) and Jevons maximum overlap star indexes, J1-J5, are listed in Table 9 below. As usual, these indexes have been normalized so that the OECD level is equal to 1 for all 7 indexes.

Table 9: Alternative Interregional PPPs for the BH Eggs

Region	P_{CPD}	$P_{GEKS\text{-}J}$	P_{J1}	P_{J2}	P_{J3}	P_{J4}	P_{J5}
1	11.9389	11.9234	11.8614	11.7133	11.8614	11.8744	12.3149
2	10.4210	10.4473	10.5529	10.4210	10.4210	10.4210	10.4210
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	3.0206	3.0199	3.0173	3.0206	3.0206	3.0206	3.0206
5	0.3637	0.3609	0.3503	0.3637	0.3637	0.3637	0.3637

Note that the star Jevons indexes P_{J2}-P_{J5} are exactly the same for regions 2-5. This follows from our earlier algebra and the fact that these 4 regions price the same products.[28]

It can be seen that the GEKS-Jevons index and the CPD index are very close for this Basic Heading Category. This is a result of the fact that almost all products are priced by the regions for this BH category. Are the resulting interregional PPPs accurate even though the number of products priced in this BH category is tiny? They may well be accurate if the regional average prices are accurate because the range of products in this category is very narrow compared to other BH categories.

Example 4: Basic Heading Category is Wine

There are 7 products in the Wine BH category of products. Here is a listing of the individual products along with their average prices by region:

Table 11: Regional Average Prices for the BH Category Wine

Product title	Region 1	Region 2	Region 3	Region 4	Region 5
Red wine, table wine, WKB	43.40769	66.44665	2.838234	24.46784	6.647364
Red wine, European, WKB	0	0	0	0	0
White wine, table wine, WKB	49.10292	65.22707	2.761142	26.1614	5.308037
Sparkling wine, WKB	61.43441	0	8.649288	46.4729	0
Red wine, Australian, WKB	0	96.41996	0	0	0
White wine, South African, WKB	55.42196	0	0	0	0
Red wine, Chilean, WKB	0	0	0	31.50088	0

No region priced product 2 and only 1 region priced products 5, 6 and 7 so we are down to only 3 useful product categories: 1, 3 and 4. Thus we have 13 positive regional product prices out of a possible 15 prices for the 3 useful wine categories. We use these 13 degrees of freedom to estimate 7 CPD parameters.

The interregional CPD, GEKS-Jevons (GEKS-J) and Jevons maximum overlap star indexes, J1-J5, are listed in Table 12 below.

Table 12: Alternative Interregional PPPs for the BH Wine

Region	P_{CPD}	$P_{GEKS\text{-}J}$	P_{J1}	P_{J2}	P_{J3}	P_{J4}	P_{J5}
1	12.4544	13.9348	12.4544	16.4918	12.4544	12.4544	16.4918
2	20.2134	21.4752	17.7597	23.5170	23.5170	19.7743	23.5170

[28] See equations (36).

3	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
4	**7.5994**	**8.1450**	**7.5994**	**9.0377**	**7.5994**	**7.5994**	**9.0377**
5	**1.8238**	**1.9377**	**1.6024**	**2.1219**	**2.1219**	**1.7842**	**2.1219**

It can be seen that the 5 Jevons star indexes differ considerably and the CPD and GEKS-Jevons indexes also exhibit a considerable amount of variation. As usual, this variation is caused by just two missing prices: if each region priced all 3 overlapping products, all 7 of the indexes listed in Table 12 would be identical. This example shows that having a small number of missing prices can sometimes lead to very uncertain PPPs.

Example 5: Basic Heading Category is Jewelry, Clocks and Watches

There are 5 products in the Jewelry, Clocks and Watches BH category of products. Here is a listing of the individual products along with their average prices by region:

Table 13: Regional Average Prices for the BH Category Jewelry, Clocks and Watches

Product title	Region 1	Region 2	Region 3	Region 4	Region 5
Wrist-watch, children's, SWATCH Flik Flak	**0**	**0**	**52.89648**	**0**	**0**
Wrist-watch, men's, CITIZEN Eco-Drive BM6060	**2947.241**	**1428.628**	**0**	**407.2255**	**50.24698**
Analog travel alarm, quartz, BL	**152.5293**	**0**	**0**	**0**	**2.718193**
Wedding ring, 14 Karat gold, BNR	**827.9919**	**0**	**75.06806**	**297.3313**	**0**
Wall clock, SEIKO	**295.5763**	**237.8775**	**0**	**0**	**8.658026**

Only one region priced product 1 so this product is dropped from the list of products. Only 12 of the remaining 20 possible regional average prices are positive. Thus the CPD regression has 12 degrees of freedom to allocate to the estimation of 8 parameters. The CPD interregional PPPs turned out to be 13.8588, 9.1675, 1.0000, 3.1523 and 0.3099 for the 5 regions.

It can be seen that regions 2 and 3 do not have any common products and regions 3 and 5 also do not have any common products. Thus a complete set of maximum overlap bilateral Jevons indexes cannot be calculated and hence the GEKS-Jevons index cannot be calculated either. It is possible to calculate maximum overlap Jevons indexes using Regions 1 and 4 as the base region but the resulting PPPs are not reliable. The products in this Basic Heading group of products are very heterogeneous and so it is extremely important that all regions price the 5 detailed products. The CPD index and the two Jevons indexes that are possible are not reliable. The missing prices in this product category have led to more or less meaningless PPPs.

It can be seen that linking the regions when there are missing prices can be a tricky business!

6. Recommendations

The interregional PPPs produced by the two methodologies discussed above are far from being satisfactory. The Diewert methodology suffers from the Sergeev critique and the limitations of the CPD methodology. Sergeev’s suggested methodology suffers from the fact that in practice, the regions may not be pricing enough products *across all regions* so that the resulting CPD interregional PPPs are not robust.

What can be done to improve the estimation of interregional PPPs for the next round of the ICP? The answer emerges from our analysis*: the key to getting more reliable BH PPPs is to price the individual products on the core list across all regions*.

For the 2017 ICP round, the regions provided 2138 regional average prices. There were 631 core list products so the maximum number of regional average prices that each region could provide is 631. Regions

1-5 provided the following number of average prices: 415, 461, 407, 415 and 440. There were 86 Basic Heading product groups. Suppose the regions could agree on a common group of products which would be priced across all regions. Suppose further that each BH group consists of 5 common products. Thus the total number of average product prices that each region would provide to the ICP is 86x5 = 430 average prices, which is close to the present number of average prices that each region provided to the ICP in 2017. Denote the average price of product n for region r in a Basic Heading group of products as P_{rn} for r = 1,…,5 and n = 1,…,5. Application of the CPD method or calculation of the Jevons star interregional indexes for each Basic Heading category would lead to interregional PPPs that are proportional to the following vector of regional average prices:

(37) $\pi \equiv [\pi_1, \pi_2,\ldots,\pi_5] \equiv \{[\Pi_{n=1}^{5} P_{1n}]^{1/5}, [\Pi_{n=1}^{5} P_{2n}]^{1/5}, \ldots, \{[\Pi_{n=1}^{5} P_{5n}]^{1/5}\}$.

Thus each component π_r of π is equal to the geometric mean of the 5 regional average prices for region r, r = 1,…,5.

The ICP may want to give some advice to the regional coordinators on how exactly to form the regional average price vectors for the 5 products in a Basic Heading group of products. If each country c in a region r could form a national average prices for product n, say p_{rcn} for r = 1,…,5; c = 1,…,C(r) and n = 1,…,5, then define the average price for product n in region r, P_{rn}, as the geometric mean of the national product n prices p_{rcn} (deflated by the within region PPPs γ_{rc}) over all countries in the region; i.e., define P_{rn} as follows:[29]

(38) $P_{rn} \equiv [\Pi_{c=1}^{C(r)} (p_{rcn}/\gamma_{rc})]^{1/C(r)}$; r = 1,…,5; n = 1,…,5.

If the regional average prices defined by (38) were used in definitions (37), then it can be seen that the 5 interregional PPPs for region r defined by (37) are proportional to the geometric mean of all (regionally deflated) product prices (for the 5 chosen products) across all countries in region r. The number of separate country product prices required to implement the PPPs defined by (37) and (38) for 86 BH groups with 5 products in each group is 86x5x146 = 62,780. This is about 53% greater than the 40,949 prices that were submitted to the ICP in 2017.

It would not be easy to implement the above suggested methodology.[30] The regional coordinators would have to agree on a product list that contained products that are in fact available in each region. But if this suggestion could be implemented, then the precision of the interregional PPPs would be much improved and the resulting interregional PPPs could be made available to the general public. Finally, the resulting methodology would be easier to explain to the public: the resulting interregional PPPs would simply be proportional to a set of Jevons indexes!

There is another possible method that could be used to harmonize the list of products that need to be priced by an individual country in order to link the regions and that is to apply some form of *similarity linking*. Selected countries in each region could be asked to price out a harmonized product list with at least one

[29] One may ask: why deflate the product n price in country c in region r, p_{rcn}, by the within region r PPP for country c, γ_{rc}; why not use the not deflated average price $P_{rn}^* \equiv [\Pi_{c=1}^{C(r)} (p_{rcn}]^{1/C(r)}$ as the appropriate regional average price for product n? The price levels (and exchange rates) for countries within region r may be very different so using P_{rn}^* instead of P_{rn} defined by (38) would give countries with high price levels undue weight in the overall average price for the region.

[30] The reality is somewhat more complicated, however, as representativity weights enter the picture. In addition, many countries cannot price all the items on the Global Core List, as consumer preferences vary internationally, often greatly, and any common list designed to be priced everywhere could be severely unrepresentative in many countries of the world.

country that is not in the selected country's region but is "most" similar in its structure of relative prices. The details for setting up such a procedure need to be worked out. This remains a topic for further research.

We have not discussed the issues raised by importance weights. Including importance weights that differed across countries would not lead to simple Jevons indexes as is the case when all products are priced and there are no importance weights. The analysis in the previous sections could be extended to deal with the case where importance weights are collected but in our opinion, importance weights are too subjective and thus it seems best to keep the methodology as simple as possible.

References

Diewert, W.E. (1995), "Axiomatic and Economic Approaches to Elementary Price Indexes", Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (2004), "On the Stochastic Approach to Linking the Regions in the ICP", Department of Economics, Discussion Paper 04-16, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z1.

Diewert, W.E. (2005), "Weighted Country Product Dummy Variable Regressions and Index Number Formulae", *Review of Income and Wealth* 51, 561-570.

Diewert, W.E. (2010), "A Note on Sergeev's Democratic Method of Linking Regions at the Basic Heading Level", Technical Advisory Group Meeting February 17-19, World Bank, Washington D.C.

Diewert, W.E. (2022), "Quality Adjustment Methods", Chapter 8 in *Consumer Price Index Theory*, Washington D.C.: International Monetary Fund, available online at: https://www.imf.org/en/Data/Statistics/cpi-manual.

Diewert, W.E. (2023), "Scanner Data, Elementary Price Indexes and the Chain Drift Problem", pp. 445-606 in *Advances in Economic Measurement*, D. Chotikapanich, A.N. Rambaldi and N. Rhode (eds.), Singapore: Palgrave Macmillan.

Gini, C. (1931), "On the Circular Test of Index Numbers", *Metron* 9:9, 3-24.

Rao, D.S. Prasada (1995), "On the Equivalence of the Generalized Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Comparisons", Working Paper No. 5, Centre for International Comparisons, University of Pennsylvania, Philadelphia.

Rao, D.S. Prasada (2005), "On the Equivalence of the Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons", *Review of Income and Wealth* 51:4, 571-580.

Sergeev, S. (2009), "The Evaluation of the Approaches Used for the Linking of the Regions in the ICP 2005", unpublished paper, Statistics Austria, December.

Sergeev, S. (2011), "Implementation of the Parallel Computations for Linking the Regions within the 2011 ICP", ICP Technical Advisory Group Meeting, World Bank, Washington D.C., October 3-4, 2011.

Sergeev, S. (2022), "Some Refinements of Global Linking Procedures", ICP Technical Advisory Group Meeting, World Bank, Washington D.C., December 7-8, 2022.

Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.

Triplett, J. E. and R. J. McDonald (1977), "Assessing the Quality Error in Output Measures: The Case of Refrigerators", *The Review of Income and Wealth* 23:2, 137-156.