OECD OCDE

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

A World Free of Poverty

THE WORLD BANK GROUP

Weighted EKS and Generalised CPD methods for Aggregation at Basic Heading Level and above Basic Heading Level

School of Economics, University of New England - Australia

Agenda item n° 12

JOINT WORLD BANK – OECD SEMINAR ON PURCHASING POWER PARITIES

Recent Advances in Methods and Applications

WASHINGTON D.C.

30 January – 2 February 2001

WEIGHTED EKS AND GENERALISED CPD METHODS FOR AGGREGATION AT BASIC HEADING LEVEL AND ABOVE BASIC HEADING LEVEL

D.S. PRASADA RAO

School of Economics
University of New England
Armidale, Australia

1 Introduction

1. The past four decades have witnessed a tremendous surge in research involving internationally comparable national income aggregates, including studies in global inequality, productivity growth and convergence and simple real per capita income comparisons. This kind of research has been made possible due to increased availability of gross domestic product (GDP) and its components across countries, expressed in a single currency unit and adjusted for price level differences across countries. Much of this development is principally under the auspices of the International Comparison Project/Program (ICP) funded jointly by a number of international organisations including the World Bank, United Nations, OECD and the European Union.

2. A major step in any international comparison exercise is the aggregation of commodity level price and quantity data. In view of the multilateral nature of the comparisons it is widely accepted that index number methods employed in such an aggregation exercise satisfy a number of properties, such as *transitivity, country symmetry, additivity* and in addition are usually expected to preserve a degree of *characteristicity*. Aggregation of price-quantity data is attempted at two levels, first is at a very detailed item level, and second stage is at a more aggregated level usually referred to as *basic heading level*.

3. The main objective of the paper is to examine in detail the Elteto-Koves-Szulc (EKS) and the country-product-dummy (CPD) methods for aggregation and propose generalisations of these methods. The EKS method has recently gained in status since it is now the preferred method of aggregation for international comparisons at the OECD and Eurostat. Given the newly acquired status, it is necessary to examine the EKS method and see if further improvements could be made. In contrast, the CPD method has long been considered to be a statistical technique that can be used in filling gaps of missing price data. In the early

phases of ICP work, the CPD method was used as a method of aggregation below the basic heading level. The CPD method, due its econometric nature, could be extended and generalised to provide a comprehensive framework for international comparisons. A major objective of the present paper is to demonstrate the versatility of the CPD method and show how different generalisations of the method could provide a viable alternative to the currently used methods for aggregation in international comparisons.

4. The outline of the paper is as follows. Section 2 describes the notation used in the paper and establishes a few concepts that are central to index number methods for international comparisons. Section 3 examines the EKS method briefly and proposes a generalised EKS system which allows for differential weights for each binary comparison reflecting the general reliability of the binary comparison. Numerical results based on alternative specification of weights are also presented. Section 4 is devoted to the CPD method and its generalisations. The section will examine various aspects of the CPD method that can provide the basis for future use of this method in international comparisons.

2 Notation and Preliminaries

5. The following notation is used throughout the paper. Let p_{ij} and q_{ij} represent, respectively, the price and quantity of i-th commodity in j-th country. If the commodities represent items below the basic heading level, quantity data are usually not available. Prices are expressed in national currency units for a specific quantity unit of the commodity. If price data are not collected or available for a given commodity in a given country, the corresponding price is considered to be missing. If the commodities represent aggregates, at or above the basic heading level, then price and quantity data are usually available. At the aggregated level, price data are usually in the form of purchasing power parities, expressed relative to a numeraire currency unit. Quantity data are derived indirectly using the value aggregates in national currencies and PPPs.

6. We also use PPP_j to denote the purchasing power parity of j-th country currency (relative to a numeraire currency) which measures the number of national currency units that have the same purchasing power as one unit of the numeraire currency unit. To facilitate discussion of the country-product-dummy (CPD) method, let P_i denote the average international price of i-th commodity. The natural logarithms of PPP_j and P_i are, respectively, denoted by π_j and η_i. Let the number of countries and commodities included in the multilateral comparison be M and N.

3 The Elteto-Koves-Szulc method

7. The EKS method, proposed by Elteto and Koves (1964) and Szulc (1964)[1], is designed to construct transitive multilateral comparisons from a matrix of binary/pairwise comparisons derived using a formula which does not satisfy the transitivity property. The EKS method in its original form uses the binary Fisher PPPs (F_{jk}: j,k=1,..M) as the starting point.

The computational form for the EKS index is given by

$$EKS_{jk} = \prod_{l=1}^{M} \left[F_{jl} \cdot F_{lk}\right]^{1/M} \tag{1}$$

[1] It is now well recognised that Gini proposed this method in 1924. We will continue to refer to this as the EKS-method as it is the case with most publications of international organisations.

where F_{jk} denotes the Fisher price index number for country k with country j as the base. Computation of F_{jk} differs depending upon the level of aggregation at which the formula is applied. For purposes of aggregation at levels above the basic heading, Fisher formula used in PPP computations is essentially the same is standard Fisher index formula, geometric mean of the Laspeyres and Paasche indices. However, the Fisher index used for aggregation below the basic heading level (where quantity or expenditure data are not available), the formula is a bit different. These aspects are discussed in further detail when generalisations of EKS are considered at levels below the basic heading level.

8. Equation (1) defines the EKS index as an unweighted geometric average of the linked (or chained) comparisons between countries j and k using each of the countries in the comparisons as a link.

9. The EKS method in (1) produces comparisons which are transitive. In addition these indices also satisfy the important least squares property that indices in (1) deviate the least from the pairwise Fisher binary comparisons.[2] This property is in line with the property of characteristicity espoused in Drechsler (1973). Since Fisher index is considered to be ideal and possesses a number of desirable properties, the EKS method has a certain appeal since it preserves the Fisher indices to the extent possible, while constructing multilateral index numbers. However, a major problem with the EKS formula is that it gives equal weights to all linked comparisons [F_{jl} . F_{lk}], effectively assuming that they are of equal reliability. Following Rao (1997) and Rao, Maddison and Lee (2000), it can be argued that in practice it is possible to show that some link comparisons are intrinsically more reliable than others. For example, in practice we find that some pairwise Fisher indices are based on price data for many commodities while in other cases comparisons are based on prices for only one or two items. It is desirable to take this information into account when constructing the EKS multilateral indices. We outline the method described in Rao (1999) and apply the new method to consider different measures of reliability.

Generalized EKS Method

10. In order to generalize the EKS method to incorporate weights to various linked comparisons involved in equation (1), it is necessary to look at the EKS method from a different angle. Suppose we wish to derive a set of index numbers I_{jk} which are transitive and minimize the log-distance from the Fisher indices, then we

$$\text{minimize} \sum_j \sum_k (\ln I_{jk} - \ln F_{jk})^2$$

$$\text{subject to } I_{jk} = I_{jl} \cdot I_{lk} \qquad \forall j,k,l$$

11. Even though this optimisation problem appears to be difficult to solve, it can be handled with considerable ease once the problem is reparametrised using the following commonly known simple result.

Result: A multilateral system of index numbers, I_{jk} (j,k=1,2,…,M), satisfy transitivity property if and only there exist M numbers $\Pi_1, \Pi_2, ..., \Pi_M$ such that, for all j and k

$$\ln I_{jk} = \Pi_k - \Pi_j$$

[2] A formal proof of this is given in Rao and Baneerjee (1984).

12. Using this result, the above problem can be restated as one of finding Π_1, Π_2,..., Π_M, which minimize

$$\sum_j \sum_k (\Pi_k - \Pi_j - \ln F_{jk})^2 \qquad (2)$$

13. Then the required index I_{jk} is defined as the ratio $\exp(\hat{\Pi}_k)/\exp(\hat{\Pi}_j)$ where (^) shows that these are solutions to the minimization problem. After some simple algebraic manipulation it can be shown that the EKS index is related to the solution above as:

$$EKS_{jk} = \frac{\exp(\hat{\Pi}_k)}{\exp(\hat{\Pi}_j)} = \exp(\hat{\Pi}_k - \hat{\Pi}_j)$$

14. Considering further equation (2), it is evident that $\hat{\Pi}$'s are the ordinary least squares estimators of Π's (which are the best linear unbiased estimators) in the following model specification

$$\begin{aligned} &\ln F_{jk} = \Pi_k - \Pi_j + u_{jk} \\ &\text{with} \quad E(u_{jk}) = 0 \quad \text{and} \quad v(u_{jk}) = \sigma^2 \end{aligned} \qquad (3)$$

15. Given the model specification in (3), it is possible to discriminate between different pairs of countries using some indicators of reliability. This can be achieves using the following model

$$\begin{aligned} &\ln \mathrm{F}_{jk} = \Pi_k - \Pi_j + \mathrm{u}_{jk} \\ &\text{with} \quad \mathrm{E}(\mathrm{u}_{jk}) = 0 \quad \text{and} \quad \mathrm{v}(\mathrm{u}_{jk}) = \frac{\sigma^2}{\mathrm{w}_{jk}} \end{aligned} \qquad (4)$$

where w_{jk} is a measure of reliability. If w_{jk} is large we consider that particular Fisher index, F_{jk}, to be reliable. Modified EKS indices can be obtained by applying generalized least squares or ordinary least squares to (4)

$$\begin{aligned} &\sqrt{w_{jk}} \ln F_{jk} = \sqrt{w_{jk}}\,\Pi_k - \sqrt{w_{jk}}\,\Pi_j + u^*_{jk} \\ &\text{with} \quad E(u^*_{jk}) = 0 \quad \text{and} \quad v(u^*_{jk}) = \sigma^2 \qquad \forall j,k = 1,..M,\ j \neq k \end{aligned} \qquad (5)$$

16. Applying least squares gives the following equations to be solved:

$$\begin{bmatrix} 2\sum_{j\neq 1}^{M} \mathrm{w}_{1j} & -2\mathrm{w}_{12} & . & . & -2\mathrm{w}_{1M} \\ -2\mathrm{w}_{21} & 2\sum_{j\neq 2}^{M} \mathrm{w}_{2j} & & & -2\mathrm{w}_{2M} \\ . & & & & . \\ . & & & & . \\ -2\mathrm{w}_{M1} & -2\mathrm{w}_{M2} & . & . & 2\sum_{j\neq M}^{M} \mathrm{w}_{Mj} \end{bmatrix} \begin{bmatrix} \hat{\Pi}_1 \\ \hat{\Pi}_2 \\ . \\ . \\ \hat{\Pi}_M \end{bmatrix} = \begin{bmatrix} -2\sum_{j\neq 1}^{M} \mathrm{w}_{1j} \ln \mathrm{F}_{1j} \\ -2\sum_{j\neq 2}^{M} \mathrm{w}_{2j} \ln \mathrm{F}_{2j} \\ . \\ . \\ -2\sum_{j\neq M}^{M} \mathrm{w}_{Mj} \ln \mathrm{F}_{Mj} \end{bmatrix} \qquad (6)$$

17. In the matrix equations, we can cancel 2 from both sides. Also, the matrix on the left-hand side is singular. So we can only solve for $\hat{\Pi}_1,..\hat{\Pi}_M$ after restricting one of the $\hat{\Pi}$'s to be zero. If we set $\hat{\Pi}_1 = 0$, this implies that $PPP_1 = \exp(\hat{\Pi}_1) = 1$ and that currency of the first country is the numeraire, in our case the US dollar.

18. Given the regression model in equation (4), it is possible to use statistical packages after defining the dependent and independent variables (in this case they are dummy variables). This step could be a little confusing and at times, depending on the dimension of the problem on hand it may be necessary to define data columns with dimensions beyond what is feasible with some packages. In such cases, it is possible to use the structure of the dummy variables to arrive at algebraic expressions which lead to considerable reductions in the dimension of the problem. In the discussion below, we provide a simple operational scheme that can be used with simple EXCEL program. The steps described below provide transparent description of the computational procedure which is easily understood by practitioners not used to matrix algebra. The following steps lead to weighted EKS index numbers which take into account measures of reliability.

Step 1 Compute the Fisher binary matrix

$$F = \begin{bmatrix} F_{11} & F_{12} & . & . & F_{1M} \\ F_{21} & F_{22} & & & F_{2M} \\ . & & & & . \\ . & & & & . \\ F_{M1} & . & . & . & F_{MM} \end{bmatrix}$$

Note that $F_{ii} = 1$ for all i and that $F_{ij} = 1/ F_{ji}$.

Step 2 Specify the weight matrix for all binary comparisons

$$w = \begin{bmatrix} w_{11} & w_{12} & . & . & w_{1M} \\ w_{21} & w_{22} & & & w_{2M} \\ . & & & & . \\ . & & & & . \\ w_{M1} & . & . & . & w_{MM} \end{bmatrix}$$

Step 3 Compute matrix on the left-hand side of equation (6), denoted by P

$$P = \begin{bmatrix} \sum_{j\neq 1}^{M} w_{1j} & -w_{12} & . & . & -w_{1M} \\ -w_{21} & \sum_{j\neq 2}^{M} w_{2j} & & & -w_{2M} \\ . & & & & . \\ . & & & & . \\ -w_{M1} & -w_{M2} & . & . & \sum_{j\neq M}^{M} w_{Mj} \end{bmatrix}$$

Step 4 Drop the first row and column, and denote the resulting matrix as P^*. Compute the inverse of P^* (P^{*-1}) using e.g. EXCEL.

Step 5 Compute the vector on the right -hand side

$$q = \begin{bmatrix} -\sum_{j\neq 1}^{M} w_{1j} \ln F_{1j} \\ -\sum_{j\neq 2}^{M} w_{2j} \ln F_{2j} \\ . \\ . \\ -\sum_{j\neq M}^{M} w_{Mj} \ln F_{Mj} \end{bmatrix}$$

Step 6 Drop the first element of q and denote the result by q^*.

Step 7 The solution for $\hat{\Pi}_2, \hat{\Pi}_3, .. \hat{\Pi}_M$, given that $\hat{\Pi}_1 = 0$, can be computed using $P^{*\,-1}$ and q^* as

$$\begin{bmatrix} \hat{\Pi}_1 \\ \hat{\Pi}_2 \\ . \\ . \\ \hat{\Pi}_M \end{bmatrix} = P^{*-1} \cdot q^*$$

Step 8 The new PPPs or modified PPPs based on weighted EKS method are given by

$$\begin{aligned} &PPP_j = \exp(\hat{\Pi}_j) \\ &with \quad PPP_1 = \exp(\hat{\Pi}_j) = 1 \end{aligned} \tag{7}$$

Weighting Schemes for the Generalized EKS System

19. Given the general structure underlying the process of according weights to different linked comparisons, it is necessary to specify the matrix weights to make the method operational. In this study we consider two sets of weights for aggregation below the branch level. These are described below.

Generalised EKS System for aggregation below the basic heading level

20. The idea of using a stochastic specification for EKS method for aggregation below the basic heading level is not new. Cuthbert and Cuthbert (1989) used a similar approach in discussing the treatment of "representative" or "characteristic" items. We refer to their result in on f the alternative models considered below.

21. Before discussing the possible application of the generalised EKS system for aggregation at levels below the basic heading, it is necessary to quickly examine the present application of the standard EKS system at this level of aggregation. Two possible scenarios concerning availability of data can be considered here. First scenario, which is not all that common, is the case where price observations are available for all the items that comprise a given basic heading in all the countries. The second case is where prices are observed for only a subset of items in each country, items priced differ across different countries. In the exposition below, we follow Ferrari and Riani (1998) and Ferrari, Gozzi and Riani (1996).

22. In each of the scenarios considered, we suggest the possible specification of the weights, w_{jk}, that can be used in computing generalised EKS indices.

i) EKS procedure when price tableau is complete

23. In this case, since quantity weights are not available, the EKS PPPs for a given basic heading are derived using a geometric mean of the prices of all the commodities included.

$$EKS_{jk} = \coprod_{i=1}^{N} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/N} \quad (8)$$

where the commodities listed within the basic heading range from i=1 to N.[3]

ii) EKS procedure when price tableau is incomplete

24. We consider the case where only N_j and N_k out of N commodities (within a given basic heading) are priced. If N^*_{jk} is the number of commodities where both countries, j and k, provide price data then the EKS indices are computed using the following Fisher binary indices.

$$F_{jk} = \prod_{i \in N_{jk}^*} \left[\frac{p_{ik}}{p_{jk}} \right]^{1/N_{jk}^*} \quad (9)$$

25. The binary index is defined as the geometric average of price relatives of commonly priced items. Since indices in (9) are not transitive, the EKS procedure, in (1), is used in generating transitive PPPs at the basic heading level.

26. An obvious point for consideration here is the link between the reliability of a binary comparison and the number of items that are commonly priced. Obviously the ideal situation is where all the items are priced. It can be postulated that reliability directly proportional to the number of items that are commonly priced, with formula in (9) breaking down where there are no commonly priced items. Therefore, it can be argued that the normal EKS procedure can be replaced by a weighted EKS procedure.

27. In this case a natural specification for w_{jk} is the reciprocal of the proportion of the number of goods in the basic headings that are priced in both countries. Thus, we suggest the use of

$$w_{jk} = \frac{N}{N^*_{jk}}$$

where N denotes the number of commodities in the basic heading.

[3] Normally a different notation is needed to denote the range of commodities within a basic heading level. Where it is obvious just a general range (i=1,2, ….,N) is used to denote the commodity list that is relevant for the computation of a particular PPP.

iii) EKS procedure with incomplete price tableau and representative goods

28. In instances where the price tableau is incomplete, some of the items that are priced in a country are likely to be representative or characteristic (and therefore important) of the consumption level in the basic heading under consideration, and the others are not representative. The Eurostat and OECD follow the following version of the EKS procedure. The Fisher index between countries j and k, for the given basic heading, is computed as a geometric average of the Laspeyres- and Paasche-type indices based on prices of representative items alone. If n_j represents the set of representative commodities in country j for which prices are also available in country k, then the Laspeyres-type index is computed as

$$L_{jk} = \prod_{i \in n_j} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/n_j}$$

and the Paascheyres-type index, based on the number of representative commodities in k which are also priced in j, is defined as

$$P_{jk} = \prod_{i \in n_k} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/n_k}$$

and the Fisher index is given by

$$F_{jk} = \sqrt{L_{jk} . P_{jk}} = \sqrt{\prod_{i \in n_j} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/n_j} \cdot \prod_{i \in n_k} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/n_k}} \qquad (10)$$

29. The resulting Fisher indices do not satisfy transitivity. The current practice at Eurostat and the OECD is to apply the EKS procedure to the matrix of Fisher-binary indices in (10). This procedure is described in OECD (1999). Ferrari and Riani (1998) examine various properties of the EKS index based on Fisher binaries in (10).

30. The quality or reliability of these indices will necessarily depend upon the number of representative items of a country for which prices are available in both countries. If n_j and n_k are both low, then the Fisher index in (10) is less reliable. Cuthbert and Cuthbert (1989, p.43) use a parametric or stochastic approach to derive the following expression for variance of $\ln F_{jk}$:

$$\text{Var} (\ln F_{jk}) = \frac{\sigma^2}{4} \left[\frac{1}{n_j} + \frac{1}{n_k} + \frac{2n_{jk}}{n_j n_k} \right] \qquad (11)$$

31. Where n_j and n_k are the number of commodities characteristic in countries j and k respectively, and n_{jk} denotes the number of goods which are characteristic in both j and k.

32. In this case the generalised EKS indices can be generated by using the following specification for w_{jk}. We use

$$w_{jk} = 1/\left(\frac{1}{4} \left[\frac{1}{n_j} + \frac{1}{n_k} + \frac{2n_{jk}}{n_j n_k} \right] \right) \qquad (12)$$

33. The use of weights specified in (12) are designed to adjust for differences in the number of characteristic products priced in different countries.

34. Given the time frame for the preparation of this paper and due to lack of published price data for goods priced in different countries, no attempt has been made to apply any of these generalisations discussed above. However, Rao and Timmer (2000) use slightly different specifications in their attempt to derive consistent multilateral comparisons of manufacturing sector purchasing power parities.

Generalised EKS System for aggregation above the basic heading level

35. Although the EKS system has been in vogue for over three decades, very little research has been devoted to modifications of the EKS system that may lead to improvements in the method and its application for aggregation above the basic heading level. Hill (1999) proposed a method, based on minimum spanning trees, of linking countries and thus replacing the EKS indices with chain based comparisons. In the discussion below, we focus on the generalisations of EKS indices using various specifications for the weights that can be used.

i) Weights based on Hill's Distance Function

36. Here we consider a measure of reliability that is based on the spread between Laspeyres and Paasche index numbers. Beginning from the work of Bortkiewicz (1924), it is generally accepted that the Laspeyres-Paasche spread reflects variability in the price and quantity ratios as well as the strength of the correlation between the price and quantity ratios over time or across countries. Van Ark, Monnikhof and Timmer (1999) provide a decomposition of the spread into the different components along these lines for many binary ICOP comparisons. Hill (1999) provides formal measures of reliability based on this spread and discusses various properties of this measure. The distance between two countries j and k (d_{jk}) is measured for all j and k by

$$d_{jk} = \left| \ln\left(\frac{L_{jk}}{P_{jk}}\right) \right| \tag{13}$$

where L_{jk} and P_{jk} may refer to price index numbers or to quantity index numbers. Since a large value of d_{jk} represents a larger spread between the Laspeyres and Paasche indices, we postulate that the weights needed for our weighted EKS method are inversely proportional to the distance function. Thus, for all j and k ($j \neq k$)

$$w_{jk} = \frac{1}{d_{jk}}$$

If only one item was matched, the weight has been put to zero.

ii) Weights based on economic distance

37. Here we consider economic distance, as measured by the relative levels of real per capita income, as a measure of reliability of the direct comparison between a pair of countries.

$$d_{jk} = \left| \ln(Y_j) - \ln(Y_k) \right| \tag{14}$$

where Y represents the real per capita income of a given country. [4] In this case, the matrix of weights are defined as:

$$w_{jk} = 1/d_{jk}$$

38. This specification implies that, if two countries are at a similar stage of development, comparison of prices between countries are likely to be similar, thus resulting in a more reliable comparison. Such comparisons are assigned higher weights in the weighted EKS procedure described above. This distance function was used in Selvanathan and Rao (1992) in generating generalised Tornqvist index numbers.

iii) Similarity in price structures

39. The preceding specifications of the weights, or distances, are largely driven by the fact that binary comparisons between countries which are dissimilar (in terms of price and/or quantity structures) are intrinsically less reliable, and, therefore, less emphasis needs to be placed on the preservation of such binary comparisons. If capturing similarity in price structure is the main purposes, it is possible to obtain measures of price similarity and use them directly in the computation of generalised EKS. Indices. Kravis et al. (1982) describe a measure of price similarity in the context of regionalisation; Szulc (1996) describes a measure of similarity based on expenditure patterns rather than prices; and van Ark, Monikhof and Timmer (1996) discuss price similarity indices using quantity weights. If s_{jk} is a measure of price similarity between two countries j and k, then we can assign w_{jk} that is inversely proportional to the similarity index, which implies that higher the value of s_{jk} implies more the weight assigned to the particular comparison within the generalised EKS framework.

40. The following similarity indices are used in the empirical illustration.

Similarity index 1: This index is drawn from Kravis et al (1982).

$$s_{jk} = \frac{\sum_{i=1}^{n} W_i \, p_{ij} \, p_{ik}}{\sqrt{\sum_{i=1}^{N} W_i \, p_{ij}^2 \; \sum_{i=1}^{N} W_i \, p_{ik}^2}} \qquad (15)$$

where p_{ij} and p_{ik} respectively denote prices of i-th commodity in countries j and k; and W_i is the weight attached to i-th commodity. Kravis et al suggest the use of a global expenditure share of i-th commodity (such a definition might require a suitable measure of global expenditure and the share of the commodity under consideration!).

Similarity index 2:

$$s_{jk}(k) = \frac{\sum_{i=1}^{N} (p_{ij}q_{ik})\;(p_{ik}q_{ik})}{\sqrt{\sum_{i=1}^{m} (p_{ij}q_{ik})^2 \; \sum_{i=1}^{m} (p_{ik}q_{ik})^2}} \quad \text{and} \quad s_{jk}(j) = \frac{\sum_{i=1}^{m} (p_{ij}q_{ij})\;(p_{ik}q_{ij})}{\sqrt{\sum_{i=1}^{m} (p_{ij}q_{ij})^2 \; \sum_{i=1}^{m} (p_{ik}q_{ij})^2}} \qquad (16)$$

4 It is obvious that there is a degree of circularity in the use of real per capita income which requires the knowledge of PPPs, which are to be estimated using the generalised EKS method. It is possible to use a two-step or iterated procedure instead of a more involved maximum likelihood procedure which can simultaneously estimate all the parameters (including the distance function) involved.

41. These similarity indices, proposed in van Ark, Monnikhof and Timmer (1996), depend upon the quantities in countries j and k. In this study, we propose to use the geometric average of $s_{jk}(j)$ and $s_{jk}(k)$ to define the weights required in the generalised EKS method.

42. Given the nature of the generalisations involved, it is possible to arrive at a number of alternative specifications for the matrix of weights. Only a few options are canvassed in this paper. Since the generalisations suggested here are based on regression framework, it would be possible to undertake specification testing to choose between alternative specifications for the weighting matrices.

Numerical Illustration

43. A number of these alternative specifications are applied to the 1993 OECD results at the basic heading level. The illustration is restricted to the good and services under the private consumption expenditure aggregate. All the prices are in the form of PPPs for the basic headings expressed using US dollar as the numeraire currency. National expenditures for different basic headings are used in deriving implicit quantities. PPPs for the private consumption expenditure, derived using Fisher, standard EKS, and generalised EKS indices based on Hill's distance function, economic distance, and the two similarity indices discussed in equations (15) and (16), are presented in Table 1.

Table 1
Purchasing Power Parities, OECD, 1993
(US dollar = numeraire)
Weighted EKS method with Alternative Distance Functions

Country	Hill (1999)	Real Income Selvanathan and Rao (1992)	Price Similarity		Fisher	EKS
			Van Ark and Trimmer (1999)	KHS (1982)		
GER	2.1585	2.1242	2.1226	2.1343	2.1448	2.1365
FRA	6.9590	6.8058	6.8576	6.8950	7.2512	6.9029
ITA	1576.31	1538.77	1554.91	1563.42	1724.86	1564.76
NLD	2.2043	2.1649	2.1776	2.1899	2.1741	2.1925
BEL	40.3695	39.4206	39.6400	39.8861	40.9384	39.8980
LUX	39.3325	37.8555	37.7763	38.0434	40.7780	38.0511
UK	0.6772	0.6602	0.6667	0.6705	0.6793	0.6710
IRE	0.6972	0.6900	0.6891	0.6932	0.7455	0.6940
DNK	9.6731	9.3942	9.5489	9.6019	9.7810	9.6156
GRC	201.95	196.61	198.79	199.96	221.65	200.07
SPA	126.45	123.34	124.58	125.38	136.56	125.41
PRT	131.99	128.06	130.24	131.04	152.64	131.10
AUT	14.6035	14.2554	14.3471	14.4331	15.7604	14.4376
SUI	2.2550	2.1965	2.2360	2.2485	2.3056	2.2507
SWE	10.6822	10.4015	10.5263	10.5851	10.9961	10.5983
FIN	6.9775	6.7828	6.8811	6.9188	7.4274	6.9268
ICE	93.7615	91.3344	92.5059	93.0192	100.3348	93.1213
NOR	9.9444	9.7183	9.7604	9.8263	10.4729	9.8316
TUR	6735.85	6579.64	6636.33	6672.74	7578.89	6680.63
AUS	1.4169	1.3828	1.3946	1.4029	1.4447	1.4036
NZL	1.6050	1.5651	1.5843	1.5928	1.6588	1.5947
JAP	202.65	196.62	200.14	201.14	220.65	201.49
CAN	1.2966	1.2881	1.2908	1.2978	1.3266	1.2991
USA	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 1 shows PPPs with USA as the base country. The Fisher and EKS parities are shown in the last column. Results based on the variants of the EKS re provided in the first four columns. Differences between the EKS and the generalised EKS, relative to Fisher binaries, are not as pronounced as expected. This may be due to the fact that the empirical application involves only OECD countries. Results in the table demonstrate the feasibility of generalising the EKS method accounting for differing levels of reliability of the binary comparisons.

4 The Country-Product-Dummy (CPD) Method

44. The CPD method represents a simple regression approach to explain levels of prices of commodities in different countries. The method postulates that the observed price of a commodity, say i-th commodity in j-th country, p_{ij}, is the product of three components: the purchasing power parity or the general price level in a country relative to other countries (denoted by PPP_j); the price level of the i-th commodity relative to other commodities (denoted by P_i) and a random disturbance term v_{ij}. The model underlying the CPD method can be stated as:

$$p_{ij} = PPP_j \, P_i \cdot v_{ij}$$

or in a logarithmic form and rewriting:

$$\begin{aligned} \ln p_{ij} &= \ln PPP_j + \ln P_i + \ln v_{ij} \\ &= \pi_j + \eta_i + u_{ij} \end{aligned} \tag{17}$$

45. Further explanation of the model and a numerical illustration can be found in Maddison and Rao (1996). In order to estimate π_j (j=1,..M) and η_i (i=1,..n), it is possible to apply ordinary least squares to the following model:

$$\ln p_{ij} = \pi_1 D_1 + \pi_2 D_2 + ... + \pi_M D_M + \eta_1 D_1^* + \eta_2 D_2^* + ... + \eta_n D_n^* + u_{ij} \tag{18}$$

where D_j's and D^*_i's are respectively country and commodity dummy variables with the property that

D_j = 1 if price observation p_{ij} belongs to country j
0 otherwise

and

Di* = 1 if price observation pij refers to i-th commodity
0 otherwise

46. From the model it is obvious that irrespective how big the data set we have, it is impossible to estimate all the parameters due to the presence of perfect multicollinearity. So it is customary to estimate all the parameters after imposing a restriction. Usually one of the parameters is set to zero. In our application of the CPD method we set $\pi_1 = 0$, or equivalently $PPP_1 = 1$. Since country 1 in our list is the United States, all the PPPs and commodity specific effects (η_i) are all estimated using US dollar as the numeraire currency.

47. While the empirical application of model (18) using a statistical package is fairly straightforward, it is quite a messy operation to create all the dummy variable. If we have 20 countries and 100 commodities, these variables will be of length 2000 and each of them needs to be constructed separately. However it is possible to apply this method using a spreadsheet program like Excel using simple matrix multiplication and inversion routines. The following approach is used in this study and is highly recommended. Using simple algebra, the normal equations underlying the least squares regressions can be shown to be of the following form

$$\left[\begin{array}{ccccc|ccccccc}
n_1 & 0 & . & . & 0 & 1 & 1 & 0 & 1 & . & . & 0 \\
0 & n_2 & . & . & 0 & 0 & 1 & 0 & 1 & . & . & . \\
 & & & & & . & & & & & & . \\
 & & & & & . & & & & & & . \\
0 & . & . & . & n_M & 1 & 0 & . & . & . & . & 1 \\
\hline
1 & 0 & . & . & 1 & m_1 & 0 & . & . & . & . & 0 \\
1 & 1 & & & 0 & 0 & m_2 & . & . & . & . & 0 \\
0 & 0 & & & . & & & & & & & \\
1 & 1 & & & . & & & & & & & \\
. & . & & & . & & & & & & & \\
. & . & & & . & & & & & & & \\
0 & . & . & . & 1 & 0 & . & . & . & . & . & m_n
\end{array}\right]
\left[\begin{array}{c}
\pi_1 \\ \pi_2 \\ . \\ . \\ \pi_M \\ \hline \eta_1 \\ \eta_2 \\ . \\ . \\ . \\ . \\ \eta_n
\end{array}\right]
=
\left[\begin{array}{c}
\sum_i \ln p_{i1} \\ \sum_i \ln p_{i2} \\ \\ \\ \sum_i \ln p_{iM} \\ \hline \sum_j \ln p_{1j} \\ \sum_j \ln p_{2j} \\ \\ \\ \\ \\ \sum_j \ln p_{nj}
\end{array}\right] \qquad (19)$$

where n_j = number of commodities for which we have price in country j and m_i = number of countries in which i-th commodity has a price.

48. Once this matrix is set up, we solved the equations after imposing the restriction equation $\pi_1 = 0$.[5] The resulting estimates $\hat{\pi}_2, \hat{\pi}_3, .., \hat{\pi}_M$ are used in obtaining the PPPs for each country using

$$PPP_j = \exp(\hat{\pi}_j) = \hat{\pi}_j^*$$

CPD Method for aggregation below the basic heading level

49. The CPD method, since the time it was first proposed by Summers (1973), has long been considered as a systematic method of filling gaps in the price tableau prior to the computation of PPPs at the basic heading level. However, the CPD method can be used directly as a method of aggregation below the basic heading level.

50. If the price tableau is complete in the first instance, then PPPs from the CPD method are equal to the geometric average of price relatives (see Ferrari et al., 1996 for a proof), which, incidentally, is the EKS index under complete price information. Thus the EKS and CPD methods for aggregation below the basic heading level provide the same PPPs, and, therefore, there is no real need for a choice between the two methods.

[5] This means dropping the first column and row of the matrix and the first elements of the vectors. The dimension of the matrix on the left-hand side might be too large for Excel to invert. In that case it is possible to use formulae for inverses of partitioned matrices as we did in this study for some branches.

51. When the price tableau is incomplete, then the CPD method produces a transitive set of PPPs taking into account all the price information[6] in a single step. Cuthbert and Cuthbert (1989) attempt to extend the standard model to incorporate any bias induced by price information pertaining to commodities that are not characteristic in a given country.

52. Given the nature of the specification of the regression model underlying the CPD method, it appears that the model has not been fully investigated in the international comparison context. In this paper we attempt to demonstrate that the CPD method can be generalised to provide a framework for computing PPPs and international prices for aggregation above the basic heading level.

Generalized CPD method for aggregation above the basic heading level

53. The CPD method has never been considered as an aggregation procedure for international comparisons even though it provides the same type of results as the G-K method. The regression estimation of the CPD model provides PPPs as well as international prices in the form of P_i (or $\exp(\eta_i)$). The principal reason for any lack of such applications is that it does not make use of any quantity or value data. Thus until recently, the CPD method has remained as an aggregation procedure below the basic heading level (where no quantity information is present) and also as a method for filling holes in price information (Summers 1973).

54. However, Rao (1995) has generalized the CPD method by making use of the quantity and value data directly into the CPD method. The basic idea behind this generalisation comes from the fact that the standard CPD regression model attempts to track the (logarithm) of the observed prices using an unweighted residual sum of squares. However in the spirit of the standard index number approach, a more appropriate procedure would be to find estimates of the parameters that are likely to track the more important commodities more closely. This is achieved by minimising a weighted residual sum of squares, with each observation weighted according to the expenditure share of the commodity in a given country. Thus the generalized CPD method suggests that estimation of equation (18)

$$\ln p_{ij} = \pi_1 D_1 + \pi_2 D_2 + \ldots + \pi_M D_M + \eta_1 D_1^* + \eta_2 D_2^* + \ldots + \eta_n D_n^* + u_{ij}$$

is conducted after weighting each observation according to its value share. This is equivalent to the application of ordinary least squares after transforming the equation premultiplied by

$$\sqrt{w_{ij}} \ln p_{ij} = \pi_1 \sqrt{w_{ij}} D_1 + \pi_2 \sqrt{w_{ij}} D_2 + \ldots + \pi_M \sqrt{w_{ij}} D_M + \eta_1 \sqrt{w_{ij}} D_1^* + \ldots + \eta_n \sqrt{w_{ij}} D_n^* + v_{ij} \qquad (20)$$

where $w_{ij} = \frac{p_{ij} q_{ij}}{\sum_{i=1}^{N} p_{ij} q_{ij}}$ is the value share of i-th commodity in j-th branch.[7]

---

6 An exception to this result is the case where price data for certain commodities are available in only one country (see Ferrari et al., 1996)

7 The approach advocated here is in contrast to the model proposed in Rao (1995) and also the philosophy underlying the stochastic approach to index number described in Clemments and Izan (1981, 1987), Selvanathan (1987, 1989) and Selvanathan and Rao (1994) where variance of the disturbance term in (18) is considered to be inversely proportional to the expenditure share of the commodity.

55. The approach underlying equation (20) is akin to the M-estimator approach followed in standard econometrics where a weighted sum of squares of residuals is minimised irrespective of the covariance matrix of the disturbances.

Equivalence of Generalised CPD and Rao System for International Comparisons

56. Though the CPD method provides estimates of PPPs and international prices, P_is, this method has never been really considered as a viable alternative to the Geary-Khamis system, which uses the same PPP and international price concepts, or the EKS system. This is partly due to the fact that the CPD method has always been considered in its unweighted formulation.

57. Rao (1995) has shown that the PPPs and international prices resulting from the application of least squares to the weighted model (20) are identical to those resulting from the Rao (1990) method for international comparisons. The Rao (1990) system is a variant of the Geary-Khamis system based on the use of log-linear equations and weights based on expenditure shares. The Rao system consistss of (M+N) log-linear equations involving M purchasing power parities PPP_j (j = 1,...,M) and N international prices P_i (i = 1,2,...,N). These are

$$PPP_j = \prod_{i=1}^{N} \left(\frac{p_{ij}}{P_i} \right)^{w_{ij}} \tag{21}$$

and

$$P_i = \prod_{j=1}^{M} \left(\frac{p_{ij}}{PPP_j} \right)^{w_{ij}^*} \tag{22}$$

where w_{ij} is the value share of i-th commodity and $w_{ij}^* = \frac{w_{ij}}{\sum_{j=1}^{M} w_{ij}}$ is the share of value share of j-th country with respect to i-th commodity. Rao (1990) provides a proof of the existence and uniqueness (up to a factor of proportionality) of the solution for the unknown parities and international prices. Among other characteristics, the weighting system employed here is invariant to the size of the country, unlike the Geary-Khamis system, since it uses essentially value shares as a basis for weighting.

58. Equivalence of the Rao-system (see Rao, 1995) and the generalised CPD method can be established using the equivalence of the normal equations from least squares method applied to (20) and equations (20) and (21) that define the Rao system.

59. This equivalence supports possible use of the CPD method, and its generalisations, for purposes of aggregation both below and above the basic heading level. In contrast to the Geary-Khamis method, the generalised CPD (and the Rao methods) are closer to Tornqvist indices due to the use of geometric averaging and value-share information. Once the additive consistency condition is relaxed, the case in favour of using generalised CPD method becomes stronger.

60. Since the CPD and its generalisations are rooted in standard regression framework, it is possible to extend the CPD model in a number of directions designed to handle various measurement issues. A few of these extensions are briefly discussed below.

61. Table 2 shows the PPPs from weighted and unweighted CPD methods and contrasts them with the Fisher and EKS PPPs. All the PPPs are computed using the 1993 OECD data at the basic heading level.

Table 2
Purchasing Power Parities for OECD Countries, 1993
(US dollar = numeraire)
Weighted Country-Product-Dummy Method

Country	Fisher	EKS	CPD	
			Unweighted	Weighted
GER	2.0483	2.1365	2.1448	2.0338
FRA	6.6292	6.9029	7.2512	6.5532
ITA	1593.43	1564.76	1724.86	1504.04
NLD	2.0834	2.1925	2.1741	2.0564
BEL	38.8305	39.8980	40.9384	37.8899
LUX	38.2795	38.0511	40.7780	35.8175
UK	0.6997	0.6710	0.6793	0.6423
IRE	0.6981	0.6940	0.7455	0.6685
DNK	9.7209	9.6156	9.7810	9.1308
GRC	203.45	200.07	221.65	188.47
SPA	130.90	125.41	136.56	118.54
PRT	133.82	131.10	152.64	129.03
AUT	14.8981	14.4376	15.7604	13.7304
SUI	2.1838	2.2507	2.3056	2.1832
SWE	10.8653	10.5983	10.9961	10.0754
FIN	7.0573	6.9268	7.4274	6.5979
ICE	95.0949	93.1213	100.3348	89.5392
NOR	9.6813	9.8316	10.4729	9.2380
TUR	6851.69	6680.63	7578.89	6321.67
AUS	1.4106	1.4036	1.4447	1.3332
NZL	1.6161	1.5947	1.6588	1.5298
JAP	185.18	201.49	220.65	187.43
CAN	1.2838	1.2991	1.3266	1.2292
USA	1.0000	1.0000	1.0000	1.0000

62. Results in Table 2 show clear differences between PPPs from weighted and unweighted CPD methods. The weighted CPD methods are comparable to the EKS and Fisher PPPs. Results from the weighted CPD appear to be generally lower than the corresponding EKS and Fisher PPPs.

CPD and the Tornqvist Index Numbers

63. The CPD technique can be used in generating Tornqvist index numbers for binary comparisons, thus providing multilateral generalisations that are identical to those proposed in Caves, Christensen and Diewert (1982). Suppose we start with the CPD specification in (17). Suppose we are interested in a binary comparison between two countries j and k. The CPD model in (17) then yields:

$$\ln p_{ij} = \pi_j + \eta_i + u_{ij}$$

$$\ln p_{ik} = \pi_k + \eta_i + u_{ik}$$

64. If we are only interested in price level comparisons involving π_j and π_k, then by taking the difference of these equations, we have

$$\ln p_{ik} - \ln p_{ij} = \pi_k - \pi_j + v_{ijk} \tag{23}$$

65. Given price data, equation (23) can be used in estimating (π_k - π_j). If a weighted least squares approach is used in estimating (π_k - π_j), with weights that are averages of expenditure shares in k and j, the resulting estimator of (π_k - π_j) equals the Tornqvist index. If equation (23) is applied to data involving all pairs of countries (j,k = 1,2,…M) then the weighted least squares estimates coincide with the CCD indices proposed in Caves et al. (1982). Proof of this result can be found in Selvanathan and Rao (1992).

Quality Adjustments and the CPD Method

66. Since the CPD method provides PPPs and international prices through a regression model, it would be possible to obtain PPPs after making allowance for quality differences. The hedonic approach to quality adjustment is based on establishing a parametric relationship between the observed price and a given vector of characteristics. If Z_1, Z_2, ..., Z_Q represent a set of quality characteristics that are deemed to be relevant in a particular empirical problem[8], and if these quality characteristics are noted for each item across all the countries, then the appropriate CPD model with quality adjustments would be:

$$\ln p_{ij} = \sum_{j=1}^{M} \pi_j D_j + \sum_{i=1}^{N} \eta_i D_i^* + \sum_{q=1}^{Q} \theta_q Z_{qij} + u_{ij} \tag{24}$$

67. Model in (24) provides estimates of π_js (and, therefore, PPP_js) after making adjustments for differences in quality of the products that are priced.

68. The use of a CPD model incorporating quality adjustments was suggested in Kokoski, Moulton and Zieschang (1996) in the context of inter-area price comparisons, and Hill, Knight and Sirmans (1997) where a model similar to (24) was used in examining housing prices over time. Actual empirical application of model (24) has significant data requirements in the form of observations on quality characteristics. Model similar to (24) is also canvassed in Tripplet (2000).

Generalised CPD Method Accounting for Spatial Autocorrelation in Price Structures

69. The CPD model and its generalisations described in equations (18 and (19) make no specific assumptions about the nature of the disturbances involved in equations. In standard linear regression models, the disturbances are assumed to have a mean equal to zero and have the same variance and are not autocorrelated. Under these assumptions least squares estimates of the parameters are the best linear unbiased estimators. Generally it is possible to test and adjust for the presence of heteroscedastic disturbances.

70. In the context of the CPD model, the possible presence of autocorrelation among disturbances across countries for a specific commodity, the presence of spatial autocorrelation, has not been investigated. Aten (1996) has demonstrated the existence of spatial autocorrelation and analysed the presence of patterns in relative price structures across geographical regions. Presence of spatially autocorrelated disturbances implies that the use of ordinary least squares no longer provides the most

[8] Though the CPD method is discussed in its full generality, it is best applied to aggregation below the basic heading level when the model is used for making quality adjustments.

efficient estimates of the parameters involved. Rao and Stevano (2001) attempts to deal with the presence of spatially autocorrelated structures with the CPD model.

71. In this paper we briefly touch upon some of the main features of the generalised CPD model with spatially autocorrelated models. Before embarking on the testing and specification issues, it is useful to find an interpretation of the disturbances in the CPD model.

From (17), we have:

$$\begin{aligned} u_{ij} &= \ln p_{ij} - \pi_j - \eta_i \\ &= \ln\left(\frac{p_{ij}/PPP_j}{P_i}\right) \end{aligned} \tag{25}$$

72. From equation (25) it is evident that the disturbance term for i-th commodity in j-th country is the logarithm of domestic price of i-th commodity, p_{ij}, converted to a common currency unit using PPP_j, expressed relative to the international price of i-th commodity. Thus the disturbance term provides a measure of price levels relative to international average prices, for each commodity in each country.

73. Rao and Stevano (2001)[9] examines all aspects of spatial autoccorelation in the context of the CPD model in (18) and obtain efficient estimates of various parameters of the CPD model after accounting for the presence of spatial autoccorelation.

74. Given the nature and general scope of the present paper, no attempt is made here to provide details of the testing procedures and econometric methods of estimation. However some selected empirical results are provided to stimulate further interest in this topic.

75. The empirical results presented make use of the 1985 global comparison results from the ICP for 56 countries with eight aggregated expenditure categories. The presence of spatial autocorrelation is likely to be more pronounced when countries from different continents and at different levels of development are included. Data for the study is drawn from United Nations (1994).

76. In order to discuss spatial autocorrelation, it is necessary to specify a matrix of spatial weights, W, where a typical element w_{jk} is a measure of spatial proximity of countries j and k. Spatial proximity does not necessarily imply physical distance, it could represent economic distance as measured by the level of trade between the two countries. However, in the following empirical illustration we simply use the physical distance between countries, as measured by the distance between capital cities, is used.[10]

77. Once the W matrix is specified, then ordinary least squares residuals are used in calculating a statistic, Moran's I-statistic, to test the presence of spatial autocorrelation. The test is conducted for different commodity groups.

[9] This paper is currently under preparation. Copies are expected to be available in April, 2001. The results presented here are preliminary and subject to changes.

[10] The author is grateful to Bettina Aten for providing data on W matrices. Moran's I-statistic is asymptotically normally distributed with a non-zero mean.

Table 3: Moran's I-Statistic
Test for the presence of Spatial Autocorrelation

Commodity	I-statistic	Z-statistic (for significance testing)
1. Food	0.50009	3.57*
2. Clothing	0.47059	3.39*
3. Rent	0.32952	2.40*
4. Furniture	0.53808	3.88*
5. Medical	0.39672	2.93*
6. Transport	0.07829	0.67
7. Education	0.57515	4.09*
8. Others	0.55644	3.99*

Note: (*) indicates significance at 5% level.

78. Results in Table 3 indicate the presence of significant spatial autocorrelation for all the commodity groups, with the exception of Transport group.

79. The next step in the process of obtaining efficient estimates of the PPPs is the estimation of spatial autocorrelation coefficients for each of the commodity groups. These are presented in Table 4. The estimated correlation coefficient is very low for transport, which is consistent with the Moran's I-statistic.

Table 4: Spatial Autocorrelation Coefficients (ρ_i)
Aggregated Commodity Groups

Commodity	Estimated value of ρ
1. Food	0.6607
2. Clothing	0.5199
3. Rent	0.7423
4. Furniture	0.5324
5. Medical	0.4069
6. Transport	0.1998
7. Education	0.5206
8. Others	0.4814

80. In Table 5, we present the purchasing power parities computed using the weighted CPD method and the maximum likelihood estimates accounting for spatial autocorrelation based on the "distance" matrix of spatial correlation weights.

Table 5
Purchasing Power Parities using Weighted CPD
Base country: United States

Country	Weighted CPD Method	
	No autocorelation	With spatial auto correlation
Germany	2.523	2.4973
France	7.179	6.8801
Italy	1254.807	1153.5172
Netherlands	2.470	2.2877
Belgium	44.512	36.5288
Luxembourg	42.619	39.4496
UK	0.559	0.4912
Ireland	0.692	0.5846
Denmark	10.055	9.1397
Greece	71.638	61.7066
Spain	88.771	77.1777
Portugal	69.140	55.6519
Austria	16.774	14.5587
Finland	6.185	5.3332
Norway	8.981	9.0290
Sweden	8.241	9.3838
Australia	1.211	1.0848
New Zealand	1.260	1.1068
Japan	211.178	185.8307
Canada	1.214	1.2818
USA	1.000	1.0000
Turkey	159.097	157.6794
Hong Kong	4.301	4.2127
Korea	430.709	420.7753
Thailand	7.238	6.6428
India	4.077	3.2916
Iran	62.831	61.2501
Sri Lanka	6.236	5.0552
Pakistan	3.866	3.6958
Philippines	5.878	5.7402
Botswana	2.142	0.4946
Egypt	0.234	0.2130
Ethiopia	0.687	0.7281
Kenya	4.037	3.8456
Malawi	0.376	0.3042
Mauritius	2.410	1.9289
Nigeria	0.766	0.6978
Sierra Leone	1.885	1.8109
Swaziland	0.516	0.4905
Tanzania	13.878	12.9100
Zambia	2.278	0.8820
Zimbabwe	0.439	0.4237
Benin	85.718	66.7808
Cameroon	127.908	111.0988
Congo	160.106	144.7286
Ivory Coast	133.409	126.4336
Madagascar	221.650	239.0097
Mali	155.009	161.1177
Morocco	2.015	1.9055
Rwanda	33.283	30.8601
Senegal	121.524	116.0764
Tunisia	28.135	0.2043
Poland	71.234	66.5788
Hungary	15.768	13.4411
Yugoslavia	93.237	93.7417
Bangladesh	6.038	5.9270

81. Results in Table 5 clearly demonstrate the feasibility of using weighted CPD method for international comparisons. The last column shows PPPs when the presence of spatial autocorrelation is accounted for. PPPs in the last column differ significantly from the standard weighted CPD results. There was a significant reduction in the standard errors (not reported here) associated with these estimates.

82. On the basis of these results it appears that significant differences in PPPs can result from different error specifications. It would be useful to examine further and check if the differences represent the effects of misspecification bias due to omission of some relevant variables. These results also indicate a flow of information emanating from the presence of significant correlation in price structures across countries, with correlation depending upon geographical distance between countries.

5 Conclusions

83. The main objective of the paper is to examine in detail the EKS and CPD methods for aggregation in the context of international comparisons, and to propose generalisations of these methods that could enhance the applicability of these methods. The paper advocated a general econometric approach to the construction of EKS indices and then extended the EKS method to assign different weights to different binary comparisons reflecting the reliability of the comparisons. The procedure outlined can be applied for levels below and above the basic heading level. The paper also examined the CPD method in considerable detail. Historically, CPD method has always been considered to be a method for handling missing data and then for aggregation below the basic heading level. Based on some recent results relating to the CPD method, the present paper argues that the generalised version of the CPD method can be very versatile in handling a range of data related issues. The fact that the CPD method uses a stochastic specification makes it possible to extend the model to incorporate quality factors and additional information based on the presence of correlation between relative price levels across countries. On the basis of the results discussed in the paper, it is clear that further research on these methods can yield significant advances in methods for aggregation used in international comparisons of prices.

References

Ark, van B., E.J. Monnikhof and M.P. Timmer (1999), ‘Prices, Quantities and Productivity in Industry: A Study of Transition Economies in a Comparative Perspective’, in R. Lipsey and A. Heston (eds), International and Interarea Comparisons of Prices, Income and Output, NBER, Chicago University Press.

Aten, B (1996), "A Mapping of International Prices", *Review of Income and Wealth.*

Bortkiewicz, L. von (1924), "Zweck und Struktur einer Preisindexzahl", Nordisk Statistik Tidskrift, vol. 3.

Carter-Hill, R., J.R. Knight and C.F. Sirmans (1997), "Estimating Capital Asset Price Indexes" , *Review of Economics and Statistics*, 226-233.

Caves, D.W., Christensen, L.R. and Diewert, W.E. (1982) "Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers", Economic Journal, 92, 73-86.

Clements, K.W. and H.Y. Izan (1981) "A Note on Estimating Divisia Index Numbers", *International Economic Review*, 22, 745-47.

Clements, K.W. and H.Y. Izan (1987) "The Measurement of Inflation: A Stochastic Approach", *Journal of Business and Economic Statistics*, 5, 339-350.

Cuthbert, J.R. and M. Cuthbert (1989), *On Aggregation Methods for Purchasing Power Parities*, Working Paper No. 56, OECD Department of Economics and Statistics, Paris.

Drechsler, L. (1973),‘Weigthing of Index Numbers in Multilateral Comparisons’, *Review of Income and Wealth*, March.

Elteto, O. and P. Koves (1964) "On an index number Computation Problem in International Comparison" (in Hungarian), *Statisztikai Szemle 42, 507-518.*

Ferrari, G. and M. Riani (1998) "On Purchasing Power Prities Calculation at the Basic Heading Level", *Statistica*, Vol LVIII, 91-108.

Ferrari, G., G. Gozzi and M. Riani (1996), "Comparing GEKS and EPD Approaches for Calculating PPPs at the Basic Heading level", in *Improving the Quality of Price Indices: CPI and PPP*, Eurostat.

Hill, R.J. (1999), "Chained PPPs and Minimum Spanning Trees" in Lipsey and Heston (eds.) *International and Interarea Comparisons of Prices, Income and Output*, NBER, Chicago University Press, pp. 327-364.

Kokoski, M.F., B.R. Moulton and K.D. Zieschang (1999) "Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation", " in Lipsey and Heston (eds.) *International and Interarea Comparisons of Prices, Income and Output*, NBER, Chicago University Press, pp. 327-364

Maddison, A. and D.S. Prasada Rao (1996), A Generalized Approach to International Comparisons of Agricultural Output and Productivity Research memorandum GD-27, Groningen: Groningen Growth and Development Centre.

OECD (1987, 1996 and 1999) Purchasing Power Parities and Real Expenditures, Department of Economics and Statistics, OECD, Paris.

Rao, D.S. Prasada (1990), "A System of Log-change Index Numbers for Multilateral Comparisons", in Salazar-Carrillo and Rao (eds.) *Comparisons of Prices and Real Product in Latin America*, North-Holland.

Rao, D.S. Prasada (1996), "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 (1997) "Aggregation Methods for International Comparison of Purchasing Power Parities and Real Income: Analytical Issues and Some Recent Developments", *Proceedings of the International Statistical Institute,* 51st Session, 197-200.

Rao, D.S. Prasada and J. Stevano (2001) "Generalised CPD method for International Comparisons Accounting for Spatial Autocorrelation in Relative Prices", Paper under preparation, University of New England.

Rao, D.S. Prasada and M. Timmer (2000) *Multilateralisation of Manufacturing Sector comparisons: Issues, Methods and Empirical Results*, Research memorandum No. GD 47, Groningen Growth and Development Centre, Groningen.

Selvanathan, E.A. (1989) "A Note on the Stochastic Approach to Index Numbers", *Journal of Economics and Business Statistics,* 7, 471-474.

Selvanathan, E.A. and D.S. Prasada Rao (1992) "An Econometric Approach to the Construction of Generalised Theirl-Tornqvist Indices for Multilateral Comparisons", *Journal of Econometrics*, 54, 335-346.

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

Szulc, B. (1964) "Index Numbers of Multilateral Regional Comparisons" (in Polish), *Przeglad Statystyczny 3, 239-254.*

Szulc, B. (1996) "Criterion for adequate Linking Paths in Chain Indices with a caste study of Multi-Country Price and Volume Comparisons", *Improving the Quality of Price Indices: CPI and PPP,* EUROSTAT Seminar December 18-20 December, 1995.

Triplett, J. (2000b) "Hedonic Valuation of 'Unpriced' Banking Services: Application to National Accounts and Consumer Price Indexes", Draft paper presented at the NBER, Summer Institute.

United Nations (1994), World Comparisons of Real Gross Domestic Product and Purchasing Power, New York.