Stochastic Approach to Computation of Purchasing Power Parities in the International Comparison Program (ICP)

D.S. Prasada Rao(*)
School of Economics
The University of Queensland

and

Gholamreza Hajargasht
Department of Economics
University of Melbourne

Abstract

The paper presents a stochastic approach based on the country-product-dummy (CPD) method to the computation of purchasing power parities (PPPs) in the International Comparison Program. The approach develops estimation strategies in conjunction with the country-product-dummy method to derive a range of multilateral index number methods for the compilation of PPPs at the basic heading level as well as at higher levels of aggregation. At the basic heading level our approach generates Jevons geometric index, arithmetic and harmonic indexes as well as the Dutot index. At higher levels of aggregation, a weighted stochastic model with alternative stochastic specifications and the method of moments (MOM) are used to derive the Geary-Khamis, Iklé, Rao and other multilateral index number methods employed in international comparisons. Expressions for computing standard errors for PPPs based on these formulae are also derived. Existence of solutions to the estimating equations derived from the weighted method of moments or the maximum weighted likelihood is also discussed. A numerical illustration based on ICP 2005 data is presented.

JEL Codes: C13; C18; C43; C80
Key words: Purchasing power parities; International Comparison Program; stochastic approach; non-additive linear models; method of moments; standard errors

Research of Rao was partly supported by ARC Discovery Grant DP0985813. Hajargasht is a Post-doctoral Fellow working on ARC Discovery Project DP1094632. The authors are grateful to the comments of two anonymous referees and to Erwin Diewert for his extensive comments and suggestions on earlier versions of this paper.

Version: July 3, 2015

(*) Corresponding author: d.rao@uq.edu.au

Introduction

Purchasing power parities (PPPs) of currencies are compiled by the International Comparison Program (ICP), administered by the World Bank under the auspices of the Statistical Commission of the United Nations and with the collaboration of the OECD, EUROSTAT and other regional organizations. PPPs are widely used by researchers and policy makers in converting national accounts data for price level and currency unit differences across countries (World Bank, 2013). The 2011 round of ICP covering 199 countries has been completed and the results are available in World Bank (2015). For example, the PPP of Indian rupee in 2011 was found to be 15.10 rupees compared to the exchange rate of 46.67 rupees for one US dollar. The methodology employed in the ICP relies on price data and national accounts expenditure data collected by the participating countries and on a range of index number formulae necessary for the aggregation process[1]. Despite the importance attached to the PPPs and results from the ICP, no standard errors or reliability measures are computed and published for PPPs and real aggregates compiled by the ICP. The main objective of the paper is to develop a stochastic approach based on the country-product-dummy (CPD) method and derive the index numbers used in the ICP using the statistical framework which in turn could be used in the computation of standard errors. As demonstrated in this paper, the new stochastic approach presented here can lead to different formulae used in the computation of PPPs at the basic heading and above the basic heading levels.

The *stochastic approach* to the construction of price index numbers has a long history with Jevons, Edgeworth, Bowley and Fisher contributing to its early development. After decades of prominence enjoyed by the *axiomatic* and *economic theoretic* approaches, stochastic approach had a resurgence through Clements and Izan (1981), Selvanathan and Rao (1994),

[1] Descriptions of these methods and their properties can be found in Rao (2013a) and Diewert (2013).

Clements, Izan and Selvanathan (2005) and Diewert (2010)[2]. These contributions consider stochastic approach to index numbers as a problem of signal extraction from the messages on price changes for different commodities over time. Diewert (2010) sketches the historical development of this approach and highlights its limitations.

A relatively new strand of stochastic approach with its roots in hedonic approach to price index number construction and the CPD model of Summers (1973) has emerged over the last two decades. Though the initial application of CPD method was for filling gaps in price data and for the computation of PPPs at the basic heading level[3], recent work of Rao (2005), Diewert (2004, 2005) and Hajargahst and Rao (2010) has demonstrated the versatility of the CPD model to generate some well-known index numbers for binary and multilateral comparisons. The main objective of our paper is to develop the CPD-based stochastic approach further through the use of modern econometric tools including the method of moments estimation, M-estimators, estimation of non-additive nonlinear models and by exploring different distributional specifications for the price observations. A related objective is to provide statistical foundations to a number of index number formulae used in the computation of PPPs which are already known to possess desirable axiomatic and economic theoretic approaches thus offering a framework for the computation of standard errors for PPPs from the ICP.

The paper is organized as follows. A brief overview of the ICP methodology is presented in Section 2. Two stages involved in the ICP are identified. The first stage combines price data from participating countries to derive PPPs at the basic heading level. In the second stage, these PPPs at the basic heading level are combined with expenditure data from national accounts to compile PPPs for GDP and its components. Section 3 outlines the CPD model based stochastic approach. Section 4 focuses on the estimation of basic heading PPPs with

[2] The discussion paper version of this paper appeared in 1995.

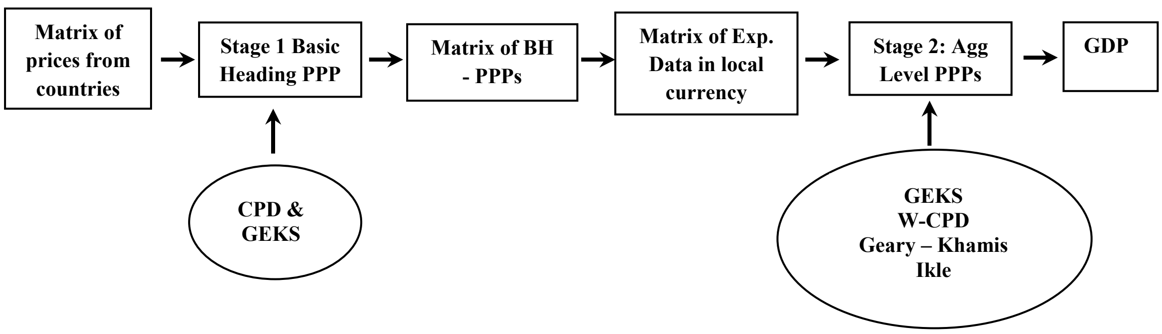
[3] Basic heading level in the ICP is the lowest level of aggregation at which expenditure data are available. This concept is further elaborated in the next section.

complete price tableau and without any weights information whereas the case of incomplete price tableau is considered in Section 5. Section 6 presents a stochastic framework based on the weighted CPD model for aggregation to levels above the basic heading and shows how standard multilateral systems due to Geary-Khamis, Iklé and Rao can be derived using this approach. A numerical illustration based on data from ICP 2005 is presented. Section 7 discusses the issue of existence of solutions to estimating equations from various models and the last section offers some concluding remarks.

1. Methodology for PPP computation in the ICP

The framework for the ICP and the process of PPP compilation is quite complex and details can be found in Rao (2013a) and World Bank (2013). The ICP uses a two-stage approach. In Figure 1, we sketch the ICP approach and list the aggregation methods relevant to PPP computation at various stages within the ICP.

Figure 1: A Schematic diagram of main steps in the ICP



In the first stage, national average prices of items belonging to a given BH are aggregated without weights using an approach similar to the construction of *elementary indexes* within the *consumer price index* (CPI). The CPD, Gini-Élteto-Köves-Szulc (GEKS) methods used in the first stage are described in Rao (2013b). The ICP at the World Bank uses the CPD method whereas the OECD-Eurostat ICP program uses the GEKS method. These are described in equations (1) and (2) below.[4] At the second stage, PPPs at the basic heading level are aggregated upwards using the available expenditure data to the GDP level or other major aggregates.[5] Diewert (2013) describes the GEKS, weighted CPD, Geary-Khamis and Iklé methods used in the second stage of ICP. The ICP at the World Bank as well as at the OECD-Eurostat uses the GEKS method whereas GK and Iklé are used when additively consistent comparisons are needed. These methods are described in column (6) of Table 1.

We consider the general case involving *M* countries and *N* commodities. Let p_{ij} represent the price of *i*-th commodity in *j*-th country expressed in local currency units (LCUs). In practice the price p_{ij} used in ICP is national average of several price quotations[6]. For the purpose of exposition in this paper, we treat price p_{ij} as a single quotation. Let PPP_j represent purchasing power parity of currency of country *j* expressed relative to a reference country currency.[7] Then PPP between currencies of any two countries *j* and *k*, PPP_{jk}, can be obtained as the ratio: PPP_k / PPP_j. The index number methods used in the ICP ensure that the resulting PPPs are *transitive* and *base invariant*.[8]

Aggregation Methods for the computation of BH PPPs: The following methods are used in aggregating price data at the BH level where no weights data are available. The main method used by the ICP at the World Bank is the *country-product-dummy* (CPD) regression model:

$$\ln p_{ij} = \eta_i + \pi_j + u_{ij} \qquad (1)$$

[4] These equations are a simplified version of the actual ICP procedure which uses a three stage aggregation procedure (over outlets and items, commodity classes and aggregates of commodity classes). For details of the weighted and unweighted Country Product Dummy method of price aggregation over items and commodity classes, see Diewert (2004) and Rao (2013b).

[5] The ICP produces and publishes results for 26 aggregates and sub-aggregates. Major aggregates of interest are Consumption, Investment, Government Expenditure and GDP.

[6] It is possible to extend the approach proposed here to take into account detailed data on price quotations or information on the number of quotations and the associated standard errors.

[7] We deliberately omit the subscript for basic heading to keep the notation simple. As aggregation below and above basic heading levels are discussed separately this should not cause any confusion.

[8] It is easy to see from the definition that $PPP_{jk} = PPP_{j\ell} \cdot PPP_{\ell k}$ for all j, k and ℓ, which is the *transitivity* requirement. Base invariance requires symmetric treatment of all the countries in the comparison. See Rao (2013a, 2013b) for further discussion on these two requirements.

where log of observed price is regressed on dummy variable for product *i* and dummy variable for country *j*. Equation (1) can be expressed in the form of a regression equation (see equation 5 below). Once the regression model (1) is estimated, PPP_j is given by $\exp(\hat{\pi}_j)$. The OECD and Eurostat use a different method (see Rao, 2013b) which is simply:[9]

$$PPP_{jk} = \prod_{i=1}^{N} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/N} \tag{2}$$

when all the items are priced in all countries and the Gini-Élteto-Köves-Szulc (GEKS) method when not all items are priced in all countries (the case of incomplete price tableau):

$$PPP_{jk}^{GEKS} = \prod_{\ell=1}^{M} \left[\prod_{i \in N_{j\ell}} \left(\frac{p_{i\ell}}{p_{ij}} \right)^{1/N_{j\ell}} \cdot \prod_{i \in N_{\ell k}} \left(\frac{p_{ik}}{p_{i\ell}} \right)^{1/N_{\ell k}} \right]^{1/M} \tag{3}$$

where N_{jk} represents the number of products that are commonly priced in countries *j* and *k*.

Aggregation Methods above the BH level: Expenditure data and therefore implicit quantity data are available for PPP computation at this level of aggregation. Diewert (2013) describes the GEKS, Iklé, Geary-Khamis and the Rao or weighted CPD methods which are commonly used in the ICP for computing PPPs for GDP and other major aggregates. Formulae for these methods, with the exception of GEKS, are given in Column (6) of Table 1.

For completeness in presentation, we describe the GEKS index. The GEKS method is a *generic* method which generates transitive indexes from a matrix of binary indexes which satisfy the country reversal test but not transitivity. Let I_{jk} represent a price index (or PPP) for country *k* with country *j* as base such that $I_{jk} \cdot I_{kj} = 1$. Then the GEKS index is given by:

$$GEKS_{jk} = \prod_{\ell=1}^{M} \left[I_{j\ell} \cdot I_{\ell k} \right]^{\frac{1}{M}}$$

The GEKS index can be implemented once the binary index number formula to compute I_{jk} is chosen. The Fisher binary index is the most commonly used index. It is possible to use

[9] This index is known as the Jevons index in CPI literature (see ILO-IMF-OECD-EUROSTAT-UNECE, 2005)

other binary indexes such as the Törnqvist index. Caves, Christensen and Diewert (1982) and Selvanathan and Rao (1994) focused on GEKS with the Törnqvist index.

In addition to these indexes, there are a number of other index number formulae used in the construction of elementary indexes (similar to BH parities) including Dutot and Carli indexes. A new system based on a weighted arithmetic index is suggested in Hajargasht and Rao (2010). Diewert (2010 and 2013) and Rao (2013b) show that all of these weighted and unweighted index number formulae can be justified in the context of the *axiomatic* or *economic theoretic* approaches to index number theory. The main objective of the paper is to show that all of these methods can be derived from the CPD model and thereby providing a unified stochastic approach to the construction of PPPs. We have been able to derive the Törnqvist-based GEKS using the stochastic approach described in this paper but not the Fisher-based GEKS method. [10]

2. Stochastic approach Based on the CPD model

The country-product-dummy (CPD) model proposed in Summers (1973) forms the basis for the stochastic approach to PPP computation presented in this paper. This method was originally proposed as a tool to fill gaps in price data but it has been in use as a method for computing *PPP*s at the basic heading level (Kravis, Heston and Summers, 1982; Diewert (2004; and Rao, 2013b). Diewert (2004) provides a useful generalization of the basic Summers (1973) model which can use data at the level of individual transactions[11]. The CPD model in its regression formulation (equation 1 above and 5 below) is commonly described as "a very simple type of hedonic regression model where the only characteristic of a commodity is the commodity itself" (Diewert, 2005, p. 561). The model in its multiplicative form is referred to

[10] The Fisher and Törnqvist indexes are known to be superlative (Diewert, 1976) and in most empirical applications these indexes are numerically quite close to each other and in fact, Diewert (1978) shows that these indexes approximate each other to the second order around an equal price and quantity point. In this sense, standard errors for the Törnqvist-based GEKS from our stochastic approach can provide useful approximation to the standard errors associated with Fisher-based GEKS indexes.

[11] The Diewert (2004) model could be potentially very useful if and when the ICP starts collecting and using individual price quotations instead of a national average price.

as the *law of one price* which postulates that the observed price, p_{ij}, of a commodity is the product of its international price, P_i, and the purchasing power parity of currency of country *j*, PPP_j. The CPD model in its multiplicative form is given by:

$$p_{ij} = P_i \cdot PPP_j \cdot u_{ij}^* \tag{4}$$

where u_{ij}^*s are random disturbance terms which are independently and identically distributed.[12] The additive form of the CPD model is obtained by taking logs on both sides:

$$\ln p_{ij} = \ln P_i + \ln PPP_j + \ln u_{ij}^* = \eta_i + \pi_j + u_{ij} = \sum_{i=1}^{N} \eta_i D_i + \sum_{j=1}^{M} \pi_j D_j^* + u_{ij} \tag{5}$$

where D_i and D_j^* are product and country dummy variables which take values 1 for commodity *i* and country *j* respectively and 0 otherwise. Equation (5) is the reason why this model is known as the country-product-dummy model. Deaton (2012, p4) considers the law of one price interpretation of the CPD model and states that "if there were no trade costs and all goods were tradable and freely traded between countries, so that the law of one price held, the residual terms in (2) would be zero".

Stochastic specifications

In order to derive estimates of the coefficients of the CPD model in (5), and the implied *PPP*s, it is necessary to specify the nature of the disturbances in models (4) and (5). In the multiplicative model (4), disturbances need to have a mean equal to 1 and are independently and identically distributed. Hajargasht and Rao (2010) use the following distributions: lognormal distribution; Gamma distribution; and inverse Gamma distribution. The specifications used are: $u_{ij}^* \sim LN(0, \sigma^2)$; $u_{ij}^* \sim Gamma(\lambda, \lambda)$; and $u_{ij}^* \sim Inverse\,Gamma(\lambda, \lambda)$.

Most of the studies including Rao (2005) and Diewert (2004; and 2005) implicitly make the assumption of lognormality of disturbances in (4) and use the additive specification of the CPD model in (5). In this paper we relax this assumption and pursue a number of estimation

[12] As demonstrated in the ensuing sections, this assumption can be easily relaxed.

strategies including the *method of maximum likelihood* which depends on the distributional assumptions and the *method of moments* which does not require any distributional assumptions but can only provide asymptotic variances. We also consider estimation of the parameters when additional information on weights associated with different commodities in different countries are also available.

Salient features of the CPD Model

The stochastic approach based on the CPD model has several advantages when applied in the ICP.

1. Hedonic characteristics: As the CPD model is a standard regression model, it is possible to introduce data on price determining characteristics such as information on outlets and rural/urban locations using suitably defined dummy variables. If available, data on quality attributes can also be incorporated. Hill and Syed (2014) use variants of the CPD model to incorporate such features in the estimation of PPPs.

2. Importance and representativity of products priced: *Comparability* of products priced across countries is considered critical for the ICP. However, products priced may not always be *representative* or *important* in the countries where they are priced. Weights reflecting importance of the products priced can be included in the CPD model[13]. Alternatively, representativeness dummy variables can be included in the model.

3. Heteroskedastic and spatially correlated disturbances: It is possible to relax the assumption of independence and identical distribution of disturbances and allow for heteroskedasticity as well as for spatial dependence in the disturbances in the CPD model. Heteroskedastic disturbances can arise when price data are in the form of averages of price quotations or when there are differences in reliability of price data across countries and commodities.

If there is spatial dependence across disturbances, then, from equation (4) we have

[13] In 2011 ICP, the general recommendation was to use 3:1 weights for products considered important.

$$u_{ij}^{*} = \frac{p_{ij}}{P_i PPP_j} = \frac{p_{ij} / PPP_j}{P_i} \sim \textit{spatially autocorrelated}$$

This means that domestic price in country *j*, converted to a common currency unit using PPPs, and expressed relative to the international average price, P_i, is correlated across countries. Spatial dependence could be due to geographical proximity or prevalence of trade.

4. CPD model and GEKS based on Törnqvist binary index numbers: From the CPD model in logarithmic form we have,

$$\ln p_{ij} = \eta_i + \pi_j + u_{ij} \quad \text{and } \ln p_{ik} = \eta_i + \pi_j + u_{ik}$$

$$\Rightarrow \ln p_{ik} - \ln p_{ij} = \ln\left(\frac{p_{ik}}{p_{ij}}\right) = \pi_k - \pi_j + v_{ij}$$

The last equation here is the same as the model in Selvanathan and Rao (1994). Using arithmetic average of expenditure shares on *i*-th commodity in countries *j* and *k* as weights, Selvanthan and Rao (1994) show that the weighted least squares of the parameters of the model in log of price changes leads to the Törnqvist-based GEKS indexes and their standard errors. This means that the stochastic approach based on the CPD model can be used in deriving Törnqvist-based GEKS indexes.

4. Basic Heading PPP Estimation using CPD Model – no weights and complete tableau

In this section we consider the computation of PPPs at the basic heading level in the first stage of the ICP as illustrated in Figure 1. We demonstrate how the simple CPD model in (5) can be used in deriving a number of well-known elementary price index numbers when there is no information on expenditures or quantities. We make a distinction between two scenarios regarding availability of price data. The first case, considered in this section, is where all the items are priced in all the countries, i.e., the case of a *complete tableau* of prices. The alternative scenario of *incomplete tableau* where not all commodities are priced in all countries is considered in Section 5. Here, disturbances in the CPD model are assumed to be independently and identically distributed. We show how four well-known formulae - the

Jevons geometric index, the arithmetic and harmonic indexes, and the Dutot index can be derived using the CPD model in a multilateral context. The Dutot index is not used as frequently as the Jevons index since it is not invariant to changes in the units of measurement and hence should only be used as an elementary index in relatively homogeneous expenditure categories. These indexes are derived using the least squares or the method of moments approaches which do not require any distributional assumptions. However, these can also be derived using the maximum likelihood method coupled with different distributional assumptions.

CPD Model – I: Jevons geometric index

We start with the CPD model in (5) which uses commodity and country dummy variables.

$$\ln p_{ij} = \ln P_i + \ln PPP_j + u_{ij} = \eta_i + \pi_j + u_{ij} = \sum_{i=1}^{N} \eta_i D_i + \sum_{j=1}^{M} \pi_j D_j^* + u_{ij} \tag{6}$$

The model can be expressed as a standard regression model. Following Rao (2005), we have

$$y_{ij} = \ln p_{ij} = \mathbf{x}_{ij}\,\boldsymbol{\beta} + v_{ij} \tag{7}$$

where $\mathbf{x}_{ij} = \left[D_1 \; D_2 \; .... \; D_N \;\; D_1^* \; D_2^* \; .... \; D_M^* \right]$ and $\boldsymbol{\beta} = \left(\eta_1 \, \eta_2 \cdots \eta_N \;\; \pi_1 \, \pi_2 \cdots \pi_M \right)'$. Stacking all the *MN* observations, the model can be written in matrix notation as: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ .The model has *MN* observations in (*M*+*N*) unknowns. We observe that the matrix **X** is of rank (*M*+*N*-*1*) and therefore parameters can be estimated only after imposing a linear restriction. Setting $\pi_M = 0$ implies, from (5), that the currency of country M is the numeraire or reference currency with $PPP_M = 1$, the remaining parameters can be estimated. Dropping the last column of **X**, we have the modified equation:

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\beta}^* + \mathbf{u} \tag{8}$$

where $\mathbf{X}^*$ is the same as matrix **X** but without the last column and $\boldsymbol{\beta}^*$ is the same as vector $\boldsymbol{\beta}$ without the last element. The least squares estimator of $\boldsymbol{\beta}^*$ and its covariance matrix are given by $\hat{\boldsymbol{\beta}}^* = \left(\mathbf{X}^{*\prime}\mathbf{X}^*\right)^{-1}\mathbf{X}^{*\prime}\mathbf{y}$ and $Var\left(\hat{\boldsymbol{\beta}}^*\right) = \sigma^2\left(\mathbf{X}^{*\prime}\mathbf{X}^*\right)^{-1}$. Using the particular structure of $\mathbf{X}^*$ whose elements are all equal to either 1 or 0, it can be shown that

$$\mathbf{X}^{*\prime}\mathbf{X}^* = \left[\begin{array}{cccc|cccc} M & 0 & . & 0 & 1 & 1 & . & 1 \\ 0 & & . & . & 1 & 1 & . & . \\ . & & & 0 & . & & & . \\ 0 & . & 0 & M & 1 & 1 & . & 1 \\ \hline 1 & 1 & . & 1 & N & 0 & . & 0 \\ 1 & 1 & & 1 & 0 & N & . & . \\ . & . & & . & . & & & 0 \\ 1 & . & . & 1 & 0 & . & 0 & N \end{array}\right] = \begin{bmatrix} M\,\mathbf{I}_N & \mathbf{J}_{N\times M-1} \\ \mathbf{J}_{M-1\times N} & N\,\mathbf{I}_{M-1} \end{bmatrix} \tag{9}$$

and its inverse matrix is given by:

$$\left(\mathbf{X}^{*\prime}\mathbf{X}^*\right)^{-1} = \begin{bmatrix} \frac{\mathbf{I}_N}{M} + \frac{(M-1)\mathbf{J}_{NxN}}{MN} & -\frac{\mathbf{J}_{NxM-1}}{N} \\ -\frac{\mathbf{J}_{M-1xN}}{N} & \frac{\mathbf{I}_{M-1} + \mathbf{J}_{M-1xM-1}}{N} \end{bmatrix} \tag{10}$$

where $\mathbf{I}_N$ is an identity matrix of size N and $\mathbf{J}_{N\times M-1}$ is a matrix of ones with size $N\times(M-1)$. The LS estimator of π_j and the resulting estimator of PPP_j can be shown to be:

$$\hat{\pi}_j = \frac{1}{N}\sum_{n=1}^{N}\left[\ln p_{nj} - \ln p_{nM}\right] \qquad and \qquad P\hat{P}P_j = \exp(\hat{\pi}_j) = \prod_{n=1}^{N}\left[\frac{p_{nj}}{p_{nM}}\right]^{1/N} \tag{11}$$

The expression for $P\hat{P}P_j$ is the geometric average of price relatives in country *j* and the reference country, *M*.[14] This is the standard form of the *Jevons geometric index.* The estimated variances of $\hat{\pi}_j$ and $P\hat{P}P_j$ are given by:

$$Est.V\left(\hat{\pi}_j\right) = \frac{2}{N}\hat{\sigma}^2 \text{ where } \hat{\sigma}^2 = \frac{\sum_{j=1}^{M}\sum_{i=1}^{N}\left(\ln p_{ij} - \hat{\eta}_i - \hat{\pi}_j\right)^2}{MN - \left(M+N-1\right)} \text{ and } EstV\left(\widehat{PPP}_j\right) \approx EstV\left(\hat{\pi}_j\right)\cdot\left(\widehat{PPP}_j\right)^2$$

We note that under assumption of lognormality of u_{ij}^* in (8), the maximum likelihood estimator of $\boldsymbol{\beta}$ is identical to the least squares estimator in (11).

CPD Models leading to Arithmetic, Harmonic and Dutot Indexes

Here we show that the *method of moments* (MOM) along with different sets of moment

[14] This result was noted by Summers (1973, p.16) which states that: “In the limiting case, when there are no missing observations in the price tableau, it can be shown that the regression procedure amounts to the computation of a set of geometric means”.

conditions can be used to estimate parameters of the CPD model leading to different indexes. For this purpose we start with the CPD model in its multiplicative form, $p_{ij} = P_i PPP_j u_{ij}^*$ with the assumption that $E(u_{ij}^*) = 1$ and $Var(u_{ij}^*) = \sigma^2$. We can transform the multiplicative CPD model in a number of different ways. For example, in the previous section we took logarithms to derive a linear model. Another transformation is to rewrite the CPD model in (4) as

$$\frac{p_{ij}}{P_i PPP_j} - 1 = u_{ij} \text{ where } u_{ij} = u_{ij}^* - 1 \quad \text{ with } E(u_{ij}) = 0 \text{ and } Var(u_{ij}) = \sigma^2 \tag{12}$$

In general, we can write the transformed CPD model as

$$r_{ij}(p_{ij}, P_i, PPP_j) = u_{ij}$$

with appropriate specification for the r_{ij} 's. This model is in the form of a non-additive non-linear regression model which can be estimated using the method of moments described in Appendix I. To estimate the unweighted models we make use of a set of moment conditions of the form

$$\frac{1}{NM}\mathbf{R}'\mathbf{r} = \mathbf{0} \tag{13}$$

where $\mathbf{r} = \{r_{ij}\}$ is a column vector of order (N×M) and $\mathbf{R}'$ is an appropriately defined $(N + M) \times (N \times M)$ matrix. Equation (13) provides a set of (*N*+M) equations in as many unknown *P*s and PPPs. The solution to the unknown parameters in (13) is referred to as the MOM estimator of the vector of unknown parameters. Table 1 summarises various choices of r_{ij} and $, \mathbf{R}$ and the resulting indexes. For the first three cases, an optimal matrix **R** has been used (see Appendix II for the derivation of **R** for the arithmetic model). Note that in all the cases considered, the resulting sets of normal equations form systems of (*M*+*N*) equations in unknown PPPs and *P*s can be solved only up to a constant of proportionality. We assume $PPP_M = 1$ which means the currency of country *M* is the reference currency. We drop the last equation which is equivalent to dropping the last row of $\mathbf{R}'$. To keep the notation simple, we continue to denote the matrix without the last row by $\mathbf{R}'$.

In Table 1 we also demonstrate how information on weights can be incorporated into MOM estimation. Let $\mathbf{W}=\{w_{ij}\}$ be a diagonal matrix with expenditure shares[15] in the diagonal. Following equations (A-10) and (A-11) in Appendix I, we can incorporate the expenditure share weights matrix $\mathbf{W}$ into MOM estimation in a straightforward manner using the weighted moment conditions:

$$\frac{1}{NM}\mathbf{R}'\mathbf{W}\mathbf{r}=0 \tag{14}$$

The weighted MOM estimators incorporating the weights matrix, $\mathbf{W}$ are shown in column (6) of Table 1. Equation (A-11) in Appendix I gives a generic form for the covariance matrix of the weighted GMM estimator.

The choice of the $\mathbf{R}$ matrix is critical in the implementation of the simple MOM in (13) and the weighted MOM in (14). It is a standard result (see e.g. Wooldridge pp 542) that the most efficient choice for $\mathbf{R}$ is $\mathbf{R}^* = \mathrm{E}\left[\partial\mathbf{r}/\partial\boldsymbol{\beta}'\right]$ where $\boldsymbol{\beta}$ is the vector of parameters of the model containing the *P*s and PPPs. In rest of the paper we denote the optimal choice of moment conditions by $\mathbf{R}^*$ and other (non-optimal) moment conditions by $\mathbf{R}$.

Table 1 represents the core of this paper and it summarizes the main results of the paper and therefore deserves clear explanation. Each row represents a variant of the CPD model specification. For example, the first row refers to the linear form of the CPD model obtained by taking logs of the CPD model in (5). The moment condition matrix that underpins the Jevons and Rao systems is shown in column (3). The resulting normal equations for the unweighted case are shown in Column (4). These equations lead to the Jevons index. Column (5) shows the expenditure share based weights matrix leading to the weighted CPD or the Rao system of equations shown in Column (6). The last column provides reference to the equations used for the computation of standard errors for the estimated PPPs. The covariance

[15] This procedure applies equally well for other types of weights. It is also not necessary that the weights in the matrix are in the form of shares adding up to 1. For example in the last row of Table1, we have quantity weights matrix that leads to the Geary-Khamis method.

matrix for the Rao system is given in equation (25) with appropriate selection of the weights matrix.

Table 1: Alternative CPD Model specifications, moment conditions, unweighted and weighted indexes and their variances

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Index	$\mathbf{r} = \{r_{ij}\}$	$\mathbf{R}'$ matrix	**Unweighted**	**Weights**	**Weighted**	**Est Variance**
Unweighted: *Jevons Geometric* ***Weighted*** *Rao System*	$r_{ij} = \ln p_{ij} - \ln P_i - \ln PPP_j$	$\mathbf{R}' = \mathbf{R}^{*\prime} = \begin{bmatrix} \mathbf{i}'_M \otimes \mathbf{I}_N \\ \mathbf{I}_M \otimes \mathbf{i}'_N \end{bmatrix}$ ***Least Squares***	$\ln \widehat{PPP}_j = \frac{1}{N}\sum_{i=1}^{N} \ln \frac{p_{ij}}{\hat{P}_i}$ $\ln \hat{P}_i = \frac{1}{M}\sum_{j=1}^{M} \ln \frac{p_{ij}}{\widehat{PPP}_j}$	$\mathbf{W}(\mathbf{w}) =$ $Diag(w_{ij})$	$\ln \widehat{PPP}_j = \sum_{i=1}^{N} w_{ij} \ln \frac{p_{ij}}{\hat{P}_i}$ $\ln \hat{P}_i = \sum_{j=1}^{M} w^*_{ij} \ln \frac{p_{ij}}{\widehat{PPP}_j}$	*Jevons* $\hat{\sigma}^2 \times$ *Eq (10)* *Rao* *Eq(25)*
Arithmetic	$r_{ij} = \frac{p_{ij}}{P_i PPP_j} - 1$	$\mathbf{R}' = \mathbf{R}^{*\prime} = Diag\begin{pmatrix} -1/\mathbf{P} \\ -1/\mathbf{PPP} \end{pmatrix} \begin{bmatrix} \mathbf{i}'_M \otimes \mathbf{I}_N \\ \mathbf{I}_M \otimes \mathbf{i}'_N \end{bmatrix}$	$\hat{P}_i = \frac{1}{M}\sum_{j=1}^{M}\left(\frac{p_{ij}}{\widehat{PPP}_j}\right)$ $\widehat{PPP}_j = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{p_{ij}}{\hat{P}_i}\right)$	$\mathbf{W}(\mathbf{w}) =$ $Diag(w_{ij})$	$\widehat{PPP}_j = \sum_{i=1}^{N} w_{ij} \frac{p_{ij}}{\hat{P}_i}$ $\hat{P}_i = \sum_{j=1}^{M} w^*_{ij} \frac{p_{ij}}{\widehat{PPP}_j}$	*Unweighted* $\hat{\sigma}^2 \times$ *Eq (15)* *Weighted* *Eq(25)*
Unweighted: *Harmonic* ***Weighted:*** *Iklé*	$r_{ij} = \frac{P_i PPP_j}{p_{ij}} - 1$	$\mathbf{R}' = \mathbf{R}^{*\prime} = Diag\begin{pmatrix} 1/\mathbf{P} \\ 1/\mathbf{PPP} \end{pmatrix} \begin{bmatrix} \mathbf{i}'_M \otimes \mathbf{I}_N \\ \mathbf{I}_M \otimes \mathbf{i}'_N \end{bmatrix}$	$\frac{1}{\widehat{PPP}_j} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{P}_i}{p_{ij}}\right)$ $\frac{1}{\hat{P}_i} = \frac{1}{M}\sum_{j=1}^{M}\frac{\widehat{PPP}_j}{p_{ij}}$	$\mathbf{W}(\mathbf{w}) =$ $Diag(w_{ij})$	$\frac{1}{\widehat{PPP}_j} = \sum_{i=1}^{N} w_{ij} \frac{\hat{P}_i}{p_{ij}}$ $\frac{1}{\hat{P}_i} = \sum_{i=1}^{M} w^*_{ij} \frac{\widehat{PPP}_j}{p_{ij}}$	*Harmonic* $\hat{\sigma}^2 \times$ *Eq (15)* *Iklé* *Eq(25)*
Unweighted *Dutot* ***Weighted*** *GK*	$r_{ij} = \frac{p_{ij}}{P_i PPP_j} - 1$	$\mathbf{R}' = Diag\begin{pmatrix} -1/\mathbf{P} \\ -1/\mathbf{PPP} \end{pmatrix} \begin{bmatrix} \mathbf{i}'_M \otimes \mathbf{I}_N \\ \mathbf{I}_M \otimes \mathbf{P}'_N \end{bmatrix}$ $\mathbf{R}^{*\prime} = Diag\begin{pmatrix} -1/\mathbf{P} \\ -1/\mathbf{PPP} \end{pmatrix} \begin{bmatrix} \mathbf{i}'_M \otimes \mathbf{I}_N \\ \mathbf{I}_M \otimes \mathbf{i}'_N \end{bmatrix}$	$\widehat{PPP}_j = \sum_{i=1}^{N} p_{ij} \Big/ \sum_{i=1}^{N} \hat{P}_i$ $\hat{P}_i = \frac{1}{M}\sum_{j=1}^{M}\left(\frac{p_{ij}}{\widehat{PPP}_j}\right)$	$\mathbf{W}(\mathbf{q}) =$ $Diag(q_{ij})$	$\widehat{PPP}_j = \sum_{i=1}^{N} p_{ij} q_{ij} \Big/ \sum_{i=1}^{N} \hat{P}_i q_{ij}$ $\hat{P}_i = \sum_{j=1}^{M}\left(\frac{p_{ij} q_{ij}}{\widehat{PPP}_j}\right) \Big/ \sum_{j=1}^{M} q_{ij}$	*Dutot* *Eq (16)* *GK* *Eq(25)*

$Diag$ represents a diagonal matrix in the arguments; $\mathbf{i}_M$ is *(Mx1)* column vector of ones; $\mathbf{I}_N$ is an identity matrix of size *N*; $\mathbf{W}(.)$ is a generic weights matrix and we specify the arguments in each case. For example, $\mathbf{W}(\mathbf{w})$ denotes a weights matrix where weights are expenditure shares.

$1/\mathbf{P} = (1/P_1 \cdots 1/P_N)'$; $1/\mathbf{PPP} = (1/PPP_1 \ \cdots \ 1/PPP_M)'$; $q_{ij} = e_{ij} / p_{ij}$; , $w_{ij} = p_{ij}q_{ij} \Big/ \sum_{i=1}^{N} p_{ij}q_{ij}$;and $w^*_{ij} = w_{ij} \Big/ \sum_{j=1}^{M} w_{ij}$.

The second row of Table 1 shows the non-additive form of the CPD model. The moment conditions in column (3) are optimal for this model. Appendix 2 shows the derivation and optimality of these moment conditions. Column (4) shows that the estimating equations are essentially in the form of arithmetic averages. The weights matrix used here is the same as that used in the case of weighted CPD shown in row (1). Column (6) shows the corresponding estimating equations which are now in the form of weighted arithmetic averages derived using the weights matrix listed in column (5).

The third row is similar to row (2) but the main difference is in the specification of the CPD equation. In column (2) corresponding to this row, the parameters of interest are in the numerator of the equation. This model uses reciprocals of what is in the second row of column (2). This specification leads to normal equations based on harmonic means as shown in columns (4) and (5), respectively for the unweighted and weighted CPD models. The moment conditions in column (3) are optimal as $\mathbf{R}^*$equals $E\left[\partial\mathbf{r}/\partial\boldsymbol{\beta}'\right]$. The weighed CPD equations shown in column (6) corresponding to row (3) and the expenditure share weights matrix in column (5) is known as the Iklé index.

A comparison of rows (1) to (3) shows how different transformations of the CPD model can lead to substantially different indexes even though the same weights matrix (shown in column 5) is used. Choice between these three methods could be guided by the stochastic properties of the disturbance term. For example, if the disturbances in the CPD model are likely to be distributed as lognormal then the model in the first row and the uneweighted arithmetic averages in column (4) and the Rao system of weighted equations in column (6) are more appropriate. An advantage of the stochastic approach proposed here is that it is non-parametric and it is possible to clearly identify the moment conditions that lead to each of the estimators. As the main objective of the paper is to demonstrate the feasibility of deriving different formulae, used in international comparisons, using the stochastic approach and, in turn, derive the standard errors we do not focus on this issue[16].

Finally, the last row uses the same specification as that used in row (2) but the moment conditions used in the case of the unweighted Dutot index are not optimal and differ from the optimal choice given in row (2). The Geary-Khamis index in the very last row of Table 1 uses the optimal choice, $\mathbf{R}*$, for moment conditions but uses the weighted MOM described in equation (14). The weighted version leads to the system of equations in column (6) which are the equations that define the Geary-Khamis index.

Now we turn to a discussion of the derivation of the variance-covariance matrices associated with the asymptotic distributions of the method of moments estimators (see Appendix I for details). Consider covariance matrices of unweighted indexes used in aggregation of item level prices to derive basic heading PPPs. As the efficient choice for $\mathbf{R}$, $\mathbf{R}*$**,** is made in the case of arithmetic and harmonic indexes, the covariance matrix for the estimated price indexes is given by (see equation A-6 in the appendix):

$$Var(\hat{\mathbf{P}},\widehat{\mathbf{PPP}}) = \hat{\sigma}^2\left[\hat{\mathbf{R}}*'\hat{\mathbf{R}}*\right]^{-1} \tag{15}$$

where $\hat{\sigma}^2 = \dfrac{\hat{\mathbf{r}}'\hat{\mathbf{r}}}{MN}$ and $\hat{\mathbf{r}}$ is a vector of residuals computed using the models specified in column 2 of Table 1 along with estimated $\hat{\mathbf{P}}$ and $\widehat{\mathbf{PPP}}$ Using the particular structure of the $\mathbf{R}*$ matrix and the formula for inverse of partitioned matrices it can be shown that

$$\left[\hat{\mathbf{R}}*'\hat{\mathbf{R}}*\right]^{-1} = Diag\begin{pmatrix}\hat{\mathbf{P}}\\ \widehat{\mathbf{PPP}}\end{pmatrix}\begin{bmatrix}\dfrac{\mathbf{I}_N}{M}+\dfrac{(M-1)\mathbf{J}_{NxN}}{MN} & -\dfrac{\mathbf{J}_{N\,xM-1}}{N}\\ -\dfrac{\mathbf{J}_{M-1xN}}{N} & \dfrac{\mathbf{I}_{M-1}+\mathbf{J}_{M-1xM-1}}{N}\end{bmatrix}Diag\begin{pmatrix}\hat{\mathbf{P}}\\ \widehat{\mathbf{PPP}}\end{pmatrix} \tag{16}$$

where $\hat{\mathbf{P}}$ and $\widehat{\mathbf{PPP}}$ are column vectors of size *N* and *M-1*. Note that the middle matrix is identical to the inverse of $\mathbf{X}^{*}{}'\mathbf{X}^{*}$ from the geometric model shown in equation (10).

[16] It is possible to devise tests similar to the standard *J-test* to choose between different sets of moment conditions. But this task is beyond the scope of this paper. Further, if the parametric approach in Hajargasht and Rao (2010) is pursued it would be possible to test the plausibility of a distributional assumption (for example, log normal) against another (like the gamma or inverse-gamma).

It can be seen that the covariance matrices for the unweighted geometric, harmonic and arithmetic indexes are the same and equal to: $Var(\hat{\mathbf{P}},\widehat{\mathbf{PPP}})=\hat{\sigma}^2\left[\hat{\mathbf{R}}*'\hat{\mathbf{R}}*\right]^{-1}$ with $\left[\hat{\mathbf{R}}*'\hat{\mathbf{R}}*\right]^{-1}$ defined by (16). Note however, that both $(\hat{\mathbf{P}},\widehat{\mathbf{PPP}})$ and $\hat{\sigma}^2$ used in (16) are different across the three models. Importantly, in geometric model $\hat{\sigma}^2$ refers to $Var(\ln u_{ij}^*)$, in arithmetic model $\hat{\sigma}^2$ refers to $Var(u_{ij}^*)$ and for harmonic model $\hat{\sigma}^2$ refers to $Var(1/u_{ij}^*)$ which are in general different[17].

It is also possible to derive the geometric, arithmetic and harmonic indexes using the maximum likelihood procedure under distributional assumptions for the disturbances. Hajargasht and Rao (2010, pp S40-42) have shown that: (i) if $u_{ij}^* \sim Gamma(\lambda,\lambda)$, the first order conditions from maximum likelihood with $Gamma(\lambda,\lambda)$ yield the arithmetic index; (ii) the harmonic index can be derived using the maximum likelihood method where u_{ij}^* has an inverse-gamma distribution; and (iii) that the lognormal specification leads to the geometric specification. These results show that with appropriate distributional assumptions we can obtain the same indexes and that their statistical properties might be more appealing in some respects [for example statistical testing of models] but here we pursue the MOM approach since it is more straightforward to incorporate covariance matrix information within this approach.

Finally, we consider the set of moment conditions that lead to a Dutot type of index. These moment conditions are not optimal, i.e. $\mathbf{R}\neq E\left[\partial\mathbf{r}/\partial\boldsymbol{\beta}'\right]=\mathbf{R}*$ with $\mathbf{P_N}$ replacing $\mathbf{I_N}$ in the optimal "moment" matrix $\mathbf{R}*$. This choice may appear arbitrary but in practice it may not be far from optimal since P_i s are numbers that are likely to be close to one in practice[18]. We consider these conditions since the weighted version of this method leads to the Geary-

[17] There is no simple relationship between the indexes. In particular, there are no arithmetic-geometric-harmonic inequality types of relationships.

[18] We note here that it there is no reason why P_i should be close to 1 though we find this to be close to unity in our empirical work reported in section 6.

Khamis index. The unweighted system leads to Dutot index for PPP_j / PPP_k. According to equation (A-7) in the appendix, with a non-optimal choice for the "moment" matrix the formula for the variance is:

$$Var(\widehat{\hat{\mathbf{P}}, \mathbf{PPP}}) = \hat{\sigma}^2 \left[\hat{\mathbf{R}}' \hat{\mathbf{R}}^* \right]^{-1} \hat{\mathbf{R}}' \hat{\mathbf{R}} \left[\hat{\mathbf{R}}^{*\prime} \hat{\mathbf{R}} \right]^{-1} \tag{17}$$

where $\hat{\mathbf{R}}^* = E\,[\partial \mathbf{r}(\mathbf{y}, \mathbf{X}, \boldsymbol{\beta}) / \partial \boldsymbol{\beta}']\big|_{\hat{\beta}}$.

5. Estimation of Basic Heading PPPs using CPD Model with Incomplete Price Tableau

In this section, we discuss how the methods discussed in Section 4 can be extended to the case when price tableau is incomplete and when not all items are priced in all the countries. Let N_j represent the number of items priced in country *j* out of a total of *N* items belonging to the basic heading; and M_i represent the number of countries in which product *i* is priced.

We first define a diagonal matrix of order $MN \times MN$ denoted by [19] $\mathbf{W}(\mathbf{d}) = Diag(d_{ij})$ with d_{ij} equal to 1 if commodity *i* is priced in country *j*, and equal to 0 otherwise. *Diag* denotes a diagonal matrix. In this case, $\mathbf{W}$ is a block diagonal matrix with one block for each country.

***CPD Model 1 – Geometric Jevons Index*:** Recall that equation (8), $\mathbf{y} = \mathbf{X}^* \boldsymbol{\beta}^* + \mathbf{u}$, was used in deriving Jevons index. This model can be modified to account for incomplete pricing by pre-multiplying both sides by $\mathbf{W}$ [20]. This gives, $\mathbf{W}\mathbf{y} = \mathbf{W}\mathbf{X}^* \boldsymbol{\beta}^* + \mathbf{W}\mathbf{u}$. Application of least squares method yields the estimator: $\left[\mathbf{X}^{*\prime} \mathbf{W}' \mathbf{X}^* \right]^{-1} \mathbf{X}^{*\prime} \mathbf{W}' \mathbf{y}$ since $\mathbf{W}$ is a diagonal matrix with zeroes and ones. We can exploit the special structure of $\mathbf{W}$ and $\mathbf{X}^*$ and show that[21]

---

[19] We use $\mathbf{W}$ as a generic weights matrix and then distinguish different versions of $\mathbf{W}$ by specifying it as a function of the type of weights used in that particular model.

[20] $\mathbf{W}$ matrix here is similar to the weights matrices in column (5) of Table 1 except that the diagonal elements here take values 0 or 1 depending on whether a particular commodity is priced in a given country.

[21] See Diewert (2004; 18) for similar equations for his model which has another layer of disaggregation (down to the level of individual transactions for each country and each item).

$$\mathbf{X}^{*\prime}\mathbf{W}'\mathbf{X}^{*} = \left[\begin{array}{cccc|cccc} M_1 & 0 & \cdots & 0 & d_{1,1} & . & . & d_{1,M-1} \\ 0 & & & \vdots & . & & & . \\ \vdots & & & 0 & . & & & . \\ 0 & \cdots & 0 & M_N & d_{N,1} & . & . & d_{N,M-1} \\ \hline d_{1,1} & . & . & d_{N,1} & N_1 & 0 & \cdots & 0 \\ . & & & . & 0 & & & \vdots \\ . & & & . & \vdots & & & 0 \\ d_{1,M-1} & . & . & d_{N,M-1} & 0 & \cdots & 0 & N_{M-1} \end{array}\right] \text{with} \left(\mathbf{X}^{*\prime}\mathbf{W}'\mathbf{X}^{*}\right)^{-1} = \begin{bmatrix} \mathbf{S}_1^{-1} & -\mathbf{A}\mathbf{S}_2^{-1} \\ -\mathbf{S}_2^{-1}\mathbf{A}' & \mathbf{S}_2^{-1} \end{bmatrix} \tag{18}$$

where $\mathbf{S_1} = Diag(M_i) - \left[\!\!\left[\left\{ \sum_{k=1}^{M-1} \frac{d_{i,k}d_{h,k}}{N_k} \right\}_{i,h} \right]\!\!\right], \mathbf{S_2} = Diag(N_j) - \left[\!\!\left[\left\{ \sum_{k=1}^{N} \frac{d_{k,j}d_{k,l}}{M_k} \right\}_{j,l} \right]\!\!\right], \mathbf{A} = \left[\!\!\left[\left\{ \frac{d_{i,j}}{M_i} \right\}_{i,j} \right]\!\!\right]$

with $i,h = 1,...,N$ and $j,l = 1,...,M-1$ where *Diag* denotes diagonal matrix.

The expression for the inverse can be used only if the matrix $\mathbf{X}^{*\prime}\mathbf{W}\mathbf{X}^{*}$ is non-singular. The condition for the feasibility of this method and for the existence of inverse for the matrix requires sufficient overlap in the items priced across countries[22]. A formal statement of the condition for the existence of solutions is given in Section 8.

***CPD Models 2, 3 and 4*:** We can estimate all indexes for the case of incomplete tableau using the following moment conditions:

$$\frac{1}{\sum_{j=1}^{N_j} N_j} \mathbf{R_w}'\mathbf{r} = \mathbf{0} \tag{19}$$

with $\mathbf{R_W}' = \mathbf{R}'\mathbf{W}$ where the $\mathbf{W}$ matrix is the diagonal matrix of zeroes and ones defined above. Using equation (A-11) with $\boldsymbol{\Omega} = \mathbf{I}$, $\mathbf{R^*} = \mathbf{R}$ and $\mathbf{WW} = \mathbf{W}$, it can be shown that for all three models, the covariance matrix of the MOM estimators is the same and it is given by

$$Var(\hat{\mathbf{P}}, \widehat{\mathbf{PPP}}) = \hat{\sigma}^2 Diag\left(\begin{array}{c}\hat{\mathbf{P}} \\ \widehat{\mathbf{PPP}}\end{array}\right)\left(\mathbf{X}^{*}\mathbf{'WX}^{*}\right)^{-1} Diag\left(\begin{array}{c}\hat{\mathbf{P}} \\ \widehat{\mathbf{PPP}}\end{array}\right) \tag{20}$$

This expression is similar to (16) except that $(\mathbf{X}^{*\prime}\mathbf{X}^{*})^{-1}$ is replaced by $(\mathbf{X}^{*\prime}\mathbf{WX}^{*})^{-1}$.

6. Weighted CPD Model for Aggregation above Basic Heading Level in the ICP

So far we have considered estimation of PPPs at basic heading level in Stage 1 of the ICP shown in Figure 1. In this section we move to Stage 2 and consider estimation of PPPs at levels of aggregation above the basic heading level. For example, we may consider PPPs for

[22] The derived expressions can also be used to show that the magnitude of the estimated standard errors for *PPP*s depends on the amount of commodity overlap across countries.

GDP, Consumption, Government and Investment. At this stage of aggregation, the ICP data consists of PPPs for 155 basic headings and expenditure in national currency units at the basic heading level for all the participating countries. Let p_{ij} represent the parity for *i*-th basic heading in *j-th* country.[23] Within the ICP, BH parity is interpreted as a price for the *composite commodity* associated with the basic heading. We recall that at the BH level, expenditure data are available. Let e_{ij} represent expenditure on *i*-th basic heading in country *j* expressed in national currency units[24]. Using this information we can define "*quantity*" as $q_{ij} = e_{ij} / p_{ij}$ and expenditure shares as: $w_{ij} = e_{ij} \Big/ \sum_{i=1}^{N} e_{ij}$. It is standard in index number methodology to use expenditure share weights to derive price index numbers. The same approach is also used in the ICP.

Let $\mathbf{W}(\mathbf{w}) = Diag\left\{w_{ij}\right\}$ be a diagonal matrix with expenditure shares in its diagonal. We can incorporate the expenditure share weights matrix **W** in the MOM estimation of non-linear additive models in a straightforward manner using the moment conditions [see equations (A-10) and (A-11) in the Appendix] using equation (14)

$$\frac{1}{NM}\mathbf{R}'\mathbf{W}\mathbf{r} = \mathbf{0}$$

Column 6 of Table 1 shows that under alternative stochastic specifications of the CPD model and under alternative estimation strategies, the Rao, arithmetic, Iklé and Geary-Khamis index formulae can be derived. These are discussed below.

Rao-system**:** Weighted least squares estimation of $\ln p_{ij} = \ln P_i + \ln PPP_j + u_{ij}$ in conjunction with expenditure shares as weights leads to the Rao system (see Rao, 2005 for a proof). The weighted least squares estimator is equivalent to the maximum likelihood estimator under the

[23] We have opted to use this notation even though in earlier sections p_{ij} was used to denote the price.

[24] We note here that e_{ij} may be zero for some basic headings in some countries. The ICP practice has been one of imputing PPPs to BHs even when there are zero expenditures.

assumption of lognormality of u_{ij}^*. Diewert (2005) derives this result for the bilateral case. In the first row of Table 1, we have listed the moment conditions (in column 2) and the weights (in column 5) that lead to the Rao-system.

Iklé System**:** The weighted MOM in (24) along with the specification of matrix **R** from equation (13) and the weights shown in the third row of Table 1 for the harmonic system lead to the Iklé system. Following Hajargasht and Rao (2010), the Iklé system can also be derived using $Inverse\, Gamma(\lambda, \lambda)$ for the disturbances in the CPD model in (5).

Arithmetic System**:** A new system based on weighted arithmetic averages was proposed in Hajargasht and Rao (2010). This system can be derived using weighted MOM in (14) along with the specification of matrix **R** shown in the second row of Table 1. These lead to the arithmetic system. Again following Hajargasht and Rao (2010), this system can be derived using $Gamma(\lambda, \lambda)$ distribution for the disturbances in the CPD model.

Geary-Khamis System: In Hajargasht and Rao (2010), it was proved that the Geary-Khamis system can be derived using weighted moment conditions with quantity weights used in constructing the **R** matrix. Using our current notation, the Geary-Khamis system can be obtained by applying (14) where **R** matrix is the one derived for the Dutot index (see third row and column 3 of Table 1) along with the diagonal matrix $\mathbf{W}(\mathbf{q}) = Diag\left\{q_{ij}\right\}$ shown in column 5 of Table 1 as weights. Diewert (2005) derives Geary-Khamis system for the case of two countries but the approach provided here is more closely integrated with the approach used in deriving the other indexes where all indexes are shown to belong to the general class of MOM estimators.

For all of these index number systems, a heteroskedastic-robust covariance matrix is given by

$$\begin{aligned} Var(\hat{\mathbf{P}}, \widehat{\mathbf{PPP}}) &= \hat{\sigma}^2 \left[\hat{\mathbf{R}}'\mathbf{W}\hat{\mathbf{R}}*\right]^{-1} \hat{\mathbf{R}}'\mathbf{W\Omega W}\hat{\mathbf{R}}\left[\hat{\mathbf{R}}*'\mathbf{W}\hat{\mathbf{R}}\right]^{-1} \text{ in general, and} \\ &= \hat{\sigma}^2 \left[\hat{\mathbf{R}}*'\mathbf{W}\hat{\mathbf{R}}*\right]^{-1} \hat{\mathbf{R}}*'\mathbf{W\Omega W}\hat{\mathbf{R}}*\left[\hat{\mathbf{R}}*'\mathbf{W}\hat{\mathbf{R}}*\right]^{-1} \text{ with optimal } \hat{\mathbf{R}}* \end{aligned} \quad (21)$$

where the $\boldsymbol{R}$ matrix varies with the system under consideration (see the appendix), $\mathbf{W}(.)$ is the weights matrix in column (5) that is appropriate for a given index number system and $\boldsymbol{\Omega}$ is the covariance matrix of $\mathbf{u}$. The heteroskedastic robust covariance matrix in (21) allows us to incorporate uncertainty of the basic heading parities used as inputs into PPP computations.

***Numerical illustration*:** The MOM approach with weighted moment conditions in (24) along with different specifications of the **R** matrix and the weights shown in column 6 of Table-1 are used in conjunction with the data on 106 basic headings belonging to Consumption aggregate within GDP for 146 countries from the 2005 ICP to compute PPPs for all countries. The following table presents PPPs computed using the four weighted indexes including Rao; arithmetic; harmonic (Iklé); and the Geary-Khamis systems for 20 selected countries.

Table **2:** PPPs using different formulae with standard errors (selected countries) Consumption aggregate, ICP 2005

Country	**Geometric Rao**		**Arithmetic**		**Harmonic Iklé**		**G-K**	
	Index	**SE**	**Index**	**SE**	**Index**	**SE**	**Index**	**SE**
Australia	1.324	0.180	1.247	0.167	1.443	0.250	1.435	0.380
Germany	0.826	0.106	0.767	0.097	0.900	0.147	0.870	0.224
UK	0.578	0.075	0.554	0.071	0.624	0.104	0.635	0.164
France	0.861	0.114	0.801	0.105	0.936	0.159	0.923	0.229
Norway	8.583	1.084	8.271	1.032	9.292	1.499	9.061	2.702
Japan	125.220	17.002	118.380	15.881	136.430	23.680	129.360	34.193
South Korea	816.090	103.500	783.720	98.208	874.020	141.700	815.000	192.230
Poland	1.800	0.211	1.711	0.198	1.985	0.297	1.841	0.474
Hungary	118.780	14.497	109.740	13.233	133.980	20.903	129.040	35.065
Russia	11.502	1.343	11.175	1.289	12.467	1.861	11.902	2.872
India	13.706	1.714	13.467	1.664	14.603	2.334	13.618	3.409
China	3.576	0.444	3.386	0.416	3.888	0.618	3.662	0.880
Indonesia	3709.100	484.900	3457.800	446.660	4103.000	685.700	3770.800	964.710
Philippines	20.359	2.617	19.509	2.478	22.420	3.684	20.031	5.415
Malaysia	1.919	0.258	1.809	0.240	2.089	0.359	2.072	0.511
South Africa	4.175	0.503	3.833	0.456	4.695	0.723	4.323	1.084
Kenya	30.251	3.848	29.680	3.730	32.298	5.252	28.467	7.668
Ghana	4333.800	553.630	4284.500	540.800	4311.200	704.030	3743.500	1004.800
Brazil	1.471	0.183	1.399	0.172	1.606	0.255	1.516	0.380
Egypt	1.852	0.231	1.792	0.221	2.000	0.319	1.869	0.476
USA	1.000		1.000		1.000		1.000	

The PPPs presented in Table 2 show the number of currency units of a country that have the same purchasing power as one US dollar with respect to the basket of goods and services included in the Consumption aggregate of GDP. For example, in the case of India 13.706 rupees have the same purchasing power as one US dollar. The main purpose of this illustration is to demonstrate the feasibility of the MOM approach and the applicability of the stochastic approach based on CPD model. The estimated PPPs for consumption - are consistent with the expectations. The general expectation is that with USA as reference country PPPs for low income countries from the Iklé and G-K methods (which are referred to as the *fixed weights indexes*), should be lower than those obtained from the weighted CPD (which is a *flexible weights* index and is considered to be a pseudo-superlative index). This is due to the fact that the Iklé and G-K methods do not allow for substitution possibilities. Of these two indexes, Iklé PPPs are expected to be closer to the Rao indexes and higher than those from G-K method as the Iklé index makes use of expenditure share weights. For some low income countries, the Iklé PPPs are higher than the PPPs from weighted geometric method. In accordance with the expectations, PPPs for higher income countries such as Australia, UK and Germany from the Iklé and G-K methods are higher than those from the Rao index. The standard errors relative to the estimated PPPs are in the range of 12 to 15 percent. These magnitudes for relative standard errors are as expected as we do not incorporate any information on possible heteroskedasticity in the variances of disturbances across countries nor any information on uncertainty regarding basic heading parities is incorporated. We note the higher relative standard errors associated with PPPs derived using the G-K method. While it is difficult to identify the source, our conjecture is that the higher standard errors reflect the variability in quantity share weights used in the G-K method. In practice, variability in quantity shares across countries is significantly higher than variability observed in expenditure share weights.

Standard Errors for other ICP Methods: Estimated covariance matrix in (27) can be used in computing variances for methods like Laspeyres, Paasche, Fisher, Tornqvist and GEKS methods used in ICP. It is possible to use the standard delta method or the bootstrap approach to compute approximate standard errors. These are not pursued in this paper.

7. Existence of solutions for the Rao, Iklé, Arithmetic and Geary-Khamis Systems

We have shown in the previous sections and in Table 1 that the Rao, Iklé, Arithmetic and G-K PPPs can be obtained as solutions to systems of simultaneous equations shown column (6) of Table 1. These methods can be used only if solutions to the unknown PPPs for these systems exist and are positive and unique up to a factor of proportionality (see Hajargasht and Rao, 2013 for details). We use graph theoretic explanation to state, without proof, the conditions under which solutions exist for these methods.

Let G denote a *graph* with countries as vertices. In our case, we have a graph with M vertices representing the countries in the comparison. Connect country j with country k using an "edge" or line if there is a commodity (a composite commodity in the case of a basic heading) that is commonly consumed in both countries. Connect two countries, j and k, if there is a commodity n such that expenditures e_{nj} and e_{nk} are both strictly positive. Construct graph G by connecting all eligible pairs of countries based on the observed expenditure data. Let this graph be known as the "*expenditure-data adjacent*" graph. A "*price-data adjacent*" graph can be similarly constructed using price observations. If prices are available only if quantities are non-zero, then the price and expenditure-data adjacent graphs would be identical. However, in the case of ICP we have no quantity or expenditure information when we consider price data below the basic heading level. In this case the price-data adjacent graph becomes relevant. At the basic heading level, the ICP practice is to impute PPPs for basic headings even when there is no expenditure attributed to the basic heading. In this case, the matrix of basic heading PPPs would be complete and the existence of solutions of the

Rao, Iklé, additive and G-K systems shown in column (6) of Table will depend on the structure of the expenditure-data adjacent graph.

A graph *G* is said to be *connected* if any pair of countries *a* and *b* can be linked through a sequence of countries that are all connected by edges.

Proposition 1: A necessary and sufficient condition for the existence of solutions to the Rao, Iklé, arithmetic and *G-K* systems of equation is that the expenditure-data graph is connected. (For proof see Hajargasht and Rao, 2013).

Proposition 2: A necessary and sufficient condition for the existence of solutions in the case of incomplete price tableau and for $\mathbf{X}^{*\prime}\mathbf{W}\mathbf{X}^{*}$ in (17) to be non-singular is that the price-data graph is connected.

In this section, we have stated easily verifiable conditions for the existence of solutions to the Rao, Iklé, arithmetic and G-K systems. Connectedness of expenditure-data or price-data graph is easy to check once the data for the ICP are collected and reported by the participating countries. The conditions stated in these two propositions are quite intuitive. They simply state that we cannot compute PPPs if the countries involved in the comparison can be partitioned into two groups of countries such that there are no items in common between these two groups of countries. In this case, the expenditure-data adjacent graph is not connected. We may note here that the results of this section are not applicable to comparisons of GDP across countries due to the negative entries for basic headings that correspond to imports and to changes in inventories. But these results hold for consumption comparisons as all the basic heading expenditures and quantities are non-negative

9. Conclusions

In this paper, we presented a stochastic approach to the computation of PPPs based on the *country-product-dummy* method and the *law of one price.* The new approach is flexible enough to provide estimates of PPPs at the basic heading level and at higher levels of aggregation. The paper showed how the stochastic approach can be used in deriving standard index number systems such as the Jevons, Dutot, arithmetic and harmonic mean indexes for use at the basic heading level as well as in deriving the Rao, Iklé, arithmetic, Geary-Khamis and the Törnqvist based GEKS indexes for aggregation above basic heading level. The proposed approach can provide standard errors for all the methods used in the ICP and is flexible enough to incorporate both uncertainty arising out of the use of national average prices (based on price surveys) and possible heteroskedasticity in disturbances for different countries. The stochastic approach discussed here offers scope to employ advanced econometric techniques leading to improved estimates for PPPs and international comparisons of prices and real incomes within the International Comparison Program at the World Bank.

REFERENCES

Cameron, A. C. and P.K. Trivedi (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, Cambridge, UK.

Caves, D.W., L.R. Christensen and W.E. Diewert (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., H.Y. Izan and E.A. Selvanathan (2005), “Stochastic Index Numbers: A Review”, Economics Discussion / Working Papers 05-08, The University of Western Australia, Department of Economics.

Deaton, A. (2012), “Calibrating Measurement Uncertainty in Purchasing Power Parity Exchange Rates”, Research Program in Development Studies, mimeographed, Princeton.

Diewert, W.E. (1976), “Exact and Superlative Index Numbers”, *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1978), “Superlative Index Numbers and Consistency in Aggregation”, *Econometrica* 46, 883-900.

Diewert, W.E. (2005), “Weighted Country Product Dummy Variable Regressions and Index Number Formulae”, *Review of Income and Wealth*, 51, 561-70.

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. (2010), “On the Stochastic Approach to Index Numbers”, pp. 235-262 in *Price and Productivity Measurement*, W.E. Diewert, Bert M. Balk, Dennis Fixler, Kevin J. Fox and Alice O. Nakamura (eds.), Trafford Press.

Diewert, W.E. (2013), “Methods of Aggregation above the Basic Heading Level within Regions”, Chapter 5 in World Bank (ed) *Measuring the Real Size of the World Economy*, World Bank, Washington DC. pp. 121-167.

Hajargasht, G and D. S. Prasada Rao (2010), “Stochastic Approach to Index Numbers for Multilateral Price Comparisons and their Standard Errors”, *Review of Income and Wealth*, Series 56, *Special Issue 1*, June, 2010, pp. S32-S58.

Hajargasht, G. and D.S. Prasada Rao (2013), “On Conditions for the Existence and Uniqueness of Index Numbers for International Price Comparisons”, *mimeographed.*

Hill, R.J. and I. A. Syed (2014) "Improving International Comparisons of Prices at Basic Heading Level: An Application to the Asia-Pacific Region," *Graz Economics Papers 2012-07*, University of Graz, Department of Economics.

ILO, IMF, OECD, UN ECE, Eurostat, and World Bank. 2004. *Consumer Price Index Manual: Theory and Practice*, (ed. Peter Hill). Geneva: International Labour Office.

Kravis, I.B., A. Heston and R. Summers (1982), *World Product and Income: International Comparisons of Real Gross Product*, The Johns Hopkins University Press, Baltimore, USA.

Rao, D.S. Prasada (1990), 'A System of Log-Change Index Numbers for Multilateral Comparisons', in *Comparisons of Prices and real Products in Latin America* (eds. Salazar-Carrillo and Rao), Contributions to Economic Analysis Series, Amsterdam: North-Holland Publishing Company.

Rao, D.S. Prasada (2005). "On the Equivalence of Weighted Country–Product–Dummy (CPD) Method and the Rao-System for Multilateral Price Comparisons." *Review of Income and Wealth* (51): 571–80.

Rao, D.S. Prasada, (2013a), "The Framework for the International Comparison Program (ICP)", Chapter 1 in World Bank (ed) *Measuring the Real Size of the World Economy*, World Bank, Washington DC. pp. 13-45.

Rao, D.S. Prasada, (2013b), "Computation of Basic Heading Purchasing Power Parities (PPPs) for Comparisons within and between Regions", Chapter 4 in World Bank (ed) *Measuring the Real Size of the World Economy*, World Bank, Washington DC. pp. 93-119.

Selvanathan, E.A. and D.S. Prasada Rao (1994), *Index Numbers: A Stochastic Approach*, Ann Arbor: The University of Michigan Press.

Summers, R. (1973). "International Price Comparisons Using Incomplete Data." *Review of Income and Wealth* (19): 1–6.

Wooldridge, J.M. (2010), *Econometric Analysis of Cross Section and Panel Data*, Second Edition, MIT Press.

World Bank (2013), *Measuring the Real Size of the World Economy* (Edited book), World Bank, Washington, DC.

Appendix I: Review of Estimation of Nonlinear Models with Non-additive Errors

Consider the nonlinear regression model

$$r(y_i, \mathbf{x}_i, \boldsymbol{\beta}) = u_i \quad i = 1, \ldots, N \tag{A-1}$$

where y_i represents the dependent variable; u_i represents the random errors; $r(y_i, \mathbf{x}_i, \boldsymbol{\beta})$ is a nonlinear function and $\mathbf{x_i}$ is a $1 \times L$ vector; $\boldsymbol{\beta}$ is a $K \times 1$ column vector of parameters; and u_i is a random disturbance term with $E(u_i) = 0$ and $Var(u_i) = \sigma^2$. We further assume that the model can be non-additive i.e. it may not be possible to write it as $y_i - g(\mathbf{x}_i, \boldsymbol{\beta}) = u_i$. Parameters of an additive model can be estimated using nonlinear least squares but in general least square criterion does not provide consistent estimators for non-additive models. Inference for these models can be based on the nonlinear GMM theory. See, for example, chapter 14 of Wooldridge (2010) or section 6.5 of Cameron and Trivedi (2005) for further information on nonlinear GMM.

Method of Moment Estimation: One approach to estimate model (A-1) is to use the method of moments by writing moment conditions such as $\mathrm{E}(\mathbf{X'u}) = \mathbf{0}$. More generally, we can base the estimation on the following *K* moment conditions:

$$\mathrm{E}(\mathbf{R(X,\boldsymbol{\beta})'u}) = \mathbf{0} \tag{A-2}$$

where ***R*** is a $N \times K$ appropriately defined vector of functions of $\mathbf{X}$ and $\boldsymbol{\beta}$. By construction there are as many moment conditions as parameters therefore a method of moment estimator can be obtained by solving the following sample moment conditions

$$\frac{1}{N}\mathbf{R(X,\hat{\boldsymbol{\beta}})'r(y,X,\hat{\boldsymbol{\beta}})} = \mathbf{0} \tag{A-3}$$

The resulting estimator is asymptotically normal with covariance matrix

$$Var(\hat{\boldsymbol{\beta}}_{MM}) = \hat{\sigma}^2 \left[\hat{\mathbf{R}}'\hat{\mathbf{R}}^*\right]^{-1} \hat{\mathbf{R}}'\hat{\mathbf{R}} \left[\hat{\mathbf{R}}^{*\prime}\hat{\mathbf{R}}\right]^{-1} \tag{A-4}$$

where $\hat{\mathbf{R}}^* = E\left(\frac{\partial \mathbf{r(y,X,\boldsymbol{\beta})}}{\partial \boldsymbol{\beta}'}\right)\bigg|_{\hat{\boldsymbol{\beta}}}$, $\mathbf{R}^*$ represents the optimal choice for the "moment" matrix

given in equation (A-5); $\hat{\mathbf{R}} = \mathbf{R(X,\hat{\boldsymbol{\beta}})}$ and $\hat{\sigma}^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{N}$ with $\hat{u}_i = r(y_i, \mathbf{x}_i, \hat{\beta})$.

One issue in the above estimation is the specification of $\mathbf{R(X,\beta)}$. Different choices for **R** might be possible however it has been shown [see e.g. Wooldridge (2010) pp 440] that the optimal choice i.e. the choice that leads to the estimator with the smallest estimated variance for the parameters is

$$\mathbf{R}^{*}(\mathbf{X},\boldsymbol{\beta}) = \mathrm{E}\left[\frac{\partial \mathbf{r}(\mathbf{y},\mathbf{X},\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} \mid \mathbf{X}\right] \tag{A-5}$$

In general, the expectation on the right hand side is not tractable unless one is willing to make strong distributional assumptions. However, for the type of models considered in this paper optimal **R** is tractable without making such assumptions. In the case of optimal choice for $\mathbf{R(X,\beta)}$, it can be shown that we have:

$$Var(\hat{\boldsymbol{\beta}}_{MM}) = \hat{\sigma}^2\left[\hat{\mathbf{R}}^{*\prime}\hat{\mathbf{R}}^{*}\right]^{-1} \tag{A-6}$$

In some of the models considered in this paper we assume a general covariance matrix for errors i.e. $Var(\mathbf{u}) = \sigma^2\boldsymbol{\Omega}$ where $\boldsymbol{\Omega}$ is known. It can be shown that in this case $\hat{\boldsymbol{\beta}}_{MM}$ obtained from (A-4) using a "moment" matrix $\mathbf{R}$ is still consistent but a robust covariance matrix can be obtained as :

$$Var(\hat{\boldsymbol{\beta}}_{MM}) = \hat{\sigma}^2\left[\hat{\mathbf{R}}'\hat{\mathbf{R}}^{*}\right]^{-1}\hat{\mathbf{R}}'\boldsymbol{\Omega}\hat{\mathbf{R}}\left[\hat{\mathbf{R}}^{*\prime}\hat{\mathbf{R}}\right]^{-1} \tag{A-7}$$

Naturally, a more efficient estimator can be obtained by incorporating $\boldsymbol{\Omega}$ into the estimation process and solving the following moment conditions

$$\frac{1}{N}\mathbf{R}(\mathbf{X},\tilde{\boldsymbol{\beta}})'\,\boldsymbol{\Omega}^{-1}\mathbf{r}(\mathbf{y},\mathbf{X},\tilde{\boldsymbol{\beta}}) = \mathbf{0} \tag{A-8}$$

with covariance matrix

$$\begin{aligned} Var(\tilde{\boldsymbol{\beta}}_{MM}) &= \hat{\sigma}^2\left[\hat{\mathbf{R}}'\boldsymbol{\Omega}^{-1}\hat{\mathbf{R}}^{*}\right]^{-1}\hat{\mathbf{R}}'\boldsymbol{\Omega}^{-1}\hat{\mathbf{R}}\left[\hat{\mathbf{R}}^{*\prime}\boldsymbol{\Omega}^{-1}\hat{\mathbf{R}}\right]^{-1} \\ &= \hat{\sigma}^2\left[\hat{\mathbf{R}}^{*\prime}\boldsymbol{\Omega}^{-1}\hat{\mathbf{R}}^{*}\right]^{-1} \text{ if optimal } \mathbf{R}^{*} \text{ is used} \end{aligned} \tag{A-9}$$

Yet, in some other models considered in the paper we have some exogenous weights w_i s that represent the importance of each observation. We can incorporate these weights into the estimation process in the following manner. Suppose $\mathbf{W} = Diag\{w_i\}$ is a diagonal $N \times N$ matrix with w_i s in its diagonal, then the moment conditions will change to

$$\frac{1}{N}\mathbf{R}(\mathbf{X},\hat{\boldsymbol{\beta}})'\mathbf{W}\mathbf{r}(\mathbf{y},\mathbf{X},\hat{\boldsymbol{\beta}}) = \mathbf{0} \qquad \text{(A-10)}$$

and the heteroskedasticity-robust covariance matrix in this case can be shown to be

$$Var(\widehat{\boldsymbol{\beta}}_{MM}) = \hat{\sigma}^2 \left[\hat{\mathbf{R}}'\mathbf{W}\hat{\mathbf{R}}^*\right]^{-1} \hat{\mathbf{R}}'\mathbf{W}\boldsymbol{\Omega}\mathbf{W}\hat{\mathbf{R}}\left[\hat{\mathbf{R}}^{*\prime}\mathbf{W}\hat{\mathbf{R}}\right]^{-1} \qquad \text{(A-11)}$$

(A-11) can be proven by using the fact that $\hat{\boldsymbol{\beta}}$ from (A-10) has the same asymptotic covariance matrix as the GMM estimator $\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \mathbf{r}'\mathbf{W}\hat{\mathbf{R}}\hat{\mathbf{R}}'\mathbf{W}\mathbf{r}$. Assuming $\mathbf{W}\hat{\mathbf{R}}\hat{\mathbf{R}}'\mathbf{W}$ as a weight matrix; $\hat{\mathbf{R}}'\mathbf{W}\hat{\mathbf{R}}^*$ is invertible and applying the standard formula for covariance matrix of a GMM estimator [see e.g. formula 6.57 pp 194 in Cameron and Trivedi (2005)] and making some simplifications gives (A-11). (A-4), (A-6) and (A-7) are special cases of (A-11). (A-9) can be obtained similarly with weight matrix $\boldsymbol{\Omega}^{-1}\mathbf{R}\mathbf{R}'\boldsymbol{\Omega}^{-1}$ [see Wooldridge (2010) pp 440 - 441 for a direct proof for this case].

Appendix II: Derivation of Optimal *R* matrix for the Arithmetic Model

Consider the CPD model $r_{ij}(p_{ij}, P_i, PPP_j) = u_{ij}$ with $\mathrm{E}(u_{ij}) = 0$

where $r_{ij} = \dfrac{p_{ij}}{P_i PPP_j} - 1$. As discussed above, we estimate the model using moment conditions such as $\dfrac{1}{NM}\mathbf{R}'\mathbf{r} = \mathbf{0}$ where the most efficient choice is $\mathbf{R}^* = \mathrm{E}\left[\dfrac{\partial \mathbf{r}}{\partial [\mathbf{P}, \mathbf{PPP}]'}\right]$. To obtain $\mathbf{R}^*$ we need to take derivative of $\mathbf{r}$ and then perform the expectation. We have

$$
\frac{\partial \mathbf{r}'}{\partial [\mathbf{P},\mathbf{PPP}]'} =
\begin{bmatrix}
\begin{matrix} -\frac{p_{11}}{P_1^2 PPP_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{p_{N1}}{P_N^2 PPP_1} \end{matrix} &
\begin{matrix} -\frac{p_{12}}{P_1^2 PPP_2} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{p_{N2}}{P_N^2 PPP_2} \end{matrix} &
\begin{matrix} \cdots \\ \\ \cdots \end{matrix} &
\begin{matrix} -\frac{p_{1M}}{P_1^2 PPP_M} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{p_{NM}}{P_N^2 PPP_M} \end{matrix} \\
\begin{matrix} -\frac{p_{11}}{P_1 PPP_1^2} & .. & -\frac{p_{N1}}{P_N PPP_1^2} \\ 0 & & \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{matrix} & & &
\begin{matrix} 0 & \cdots & 0 \\ & \ddots & \vdots \\ & & 0 \\ -\frac{p_{1M}}{P_1 PPP_{1M}^2} & \cdots & -\frac{p_{NM}}{P_N PPP_M^2} \end{matrix}
\end{bmatrix}
$$

Using the fact that $E\left[\frac{p_{ij}}{P_i PPP_j}\right] = 1$, we have

$$
\mathbf{R}^{*\prime} = E\left\{\frac{\partial \mathbf{r}'}{\partial [\mathbf{P},\mathbf{PPP}]'}\right\} =
\begin{bmatrix}
\begin{matrix} -\frac{1}{P_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{1}{P_N} \end{matrix} &
\begin{matrix} -\frac{1}{P_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{1}{P_N} \end{matrix} &
\begin{matrix} \cdots \\ \\ \cdots \end{matrix} &
\begin{matrix} -\frac{1}{P_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{1}{P_N} \end{matrix} \\
\begin{matrix} -\frac{1}{PPP_1} & \cdots & -\frac{1}{PPP_M} \\ 0 & & \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{matrix} & & &
\begin{matrix} 0 & \cdots & 0 \\ & \ddots & \vdots \\ & & 0 \\ -\frac{1}{PPP_1} & \cdots & -\frac{1}{PPP_M} \end{matrix}
\end{bmatrix}
= Diag\begin{pmatrix} -1/\mathbf{P} \\ -1/\mathbf{PPP} \end{pmatrix} \begin{bmatrix} \mathbf{i}_M' \otimes \mathbf{I}_N \\ \mathbf{I}_M \otimes \mathbf{i}_N' \end{bmatrix}
$$

Note that this optimal choice for **R** is shown in column 2 of Table 1 in the paper.

Appendix III: Standard Error for Dutot Index with Complete Tableau

To obtain the covariance matrix for Dutot index we employ formula $\hat{\sigma}^2\left[\hat{\mathbf{R}}'\hat{\mathbf{R}}^*\right]^{-1}\hat{\mathbf{R}}'\hat{\mathbf{R}}\left[\hat{\mathbf{R}}^{*\prime}\hat{\mathbf{R}}\right]^{-1}$. Using inverse of partitioned matrices and some tedious algebraic manipulations, the following formula for covariance matrix of the estimates can be obtained

$$VAR\left([\hat{\mathbf{P}};\widehat{\mathbf{PPP}}]\right) = \hat{\sigma}^2 Diag\begin{pmatrix}\hat{\mathbf{P}}\\ \widehat{\mathbf{PPP}}\end{pmatrix}\begin{bmatrix}\frac{\mathbf{I}_N}{M}+\frac{(M-1)\mathbf{J}_{NxN}}{MA} & -\frac{B\mathbf{J}_{N\,xM-1}}{A}\\ -\frac{B\mathbf{J}_{M-1xN}}{A} & \frac{\mathbf{I}_{M-1}}{C}+\frac{N}{A}\left(\frac{1}{C}+\frac{(1-B)}{\sum_{i=1}^{N}\hat{P}_i}\right)\mathbf{J}_{M-1xM-1}\end{bmatrix} Diag\begin{pmatrix}\hat{\mathbf{P}}\\ \widehat{\mathbf{PPP}}\end{pmatrix}$$

where $A = N + M\sum_{i=1}^{N}\left(\hat{P}_i - 1\right)$, $B = 1 - \frac{M}{H}\sum_{i=1}^{N}\left(\hat{P}_i^2 - \hat{P}_i\right)$ and $C = \left(\sum_{i=1}^{N}\hat{P}_i\right)^2 \Big/ \sum_{i=1}^{N}\hat{P}_i^2$. Note that if $\hat{P}_i$ s are numbers close to one, then A, B and C are expected to be close to N, 1 and N respectively. Therefore the estimated covariance matrix should be close to the one from arithmetic index.