Review of Income and Wealth
Series 56, Special Issue 1, June 2010

STOCHASTIC APPROACH TO INDEX NUMBERS FOR MULTILATERAL PRICE COMPARISONS AND THEIR STANDARD ERRORS

by Gholamreza Hajargasht

Shiraz University, Iran and CEPA, The University of Queensland

and

D. S. Prasada Rao*

CEPA, The University of Queensland

The main objective of the paper is to demonstrate that a number of widely used multilateral index numbers for international comparisons of purchasing power parities (PPPs) and real incomes can be derived using the stochastic approach. The paper shows that price index numbers from commonly used methods like the Iklé, the Rao-weighted, and an additive multilateral system are all estimators of the parameters of the country–product–dummy (CPD) model. The advantage of the stochastic approach is that we can derive standard errors for the estimates of the purchasing power parities (PPPs). The PPPs and the parameters of the stochastic model are estimated using a weighted maximum likelihood procedure under different stochastic specifications for the disturbance term. Estimates of PPPs and their standard errors for OECD countries using the proposed methods are presented. The paper also outlines a method of moments approach to the estimation of PPPs under the stochastic approach. The paper shows how the Geary–Khamis system of multilateral index numbers is a method of moments estimator of the parameters of the CPD model. The paper therefore provides a coherent stochastic framework for the Geary–Khamis system and derives standard errors of the Geary–Khamis PPPs.

1. Introduction

International comparisons of real income, consumption, investment, and other national income aggregates rely on purchasing power parities (PPP) compiled under the auspices of the International Comparison Program (ICP) conducted by international organizations including the World Bank, OECD, EUROSTAT, and the United Nations. Purchasing power parities are computed using price data collected from the participating countries. PPP compilation within the ICP is undertaken at two levels—at the basic heading level and at a more

Note: The authors gratefully acknowledge support received through NSF Grant No. 27-3457-00-0-79-195 for the project entitled "Integrating Expenditure and Production Estimates in International Comparisons." The authors are grateful to Bert Balk and Erwin Diewert for their comments on contents of the paper. The revised version of the paper has greatly benefited from the comments of two anonymous referees. A part of this work was undertaken when Gholamreza Hajargasht was a post-doctoral fellow at the Centre for Efficiency and Productivity Analysis, at the University of Queensland. Recent work by D. S. Prasada Rao was supported by a grant from the Australian Research Council in the form of an Australian Professorial Fellowship associated with ARC Discovery Project DP0985813.

*Correspondence to: D. S. Prasada Rao, Centre for Efficiency and Productivity Analysis (CEPA), School of Economics, The University of Queensland, Brisbane, Australia (p.rao@economics.uq.edu.au).

aggregated level.[1] At the basic heading level price data are aggregated without any weights to yield PPPs for various basic headings. The basic heading PPPs are then aggregated to yield PPPs for higher level aggregates like consumption, investment, and gross domestic product. The main focus of the paper is on the second step involving the aggregation above the basic heading level where weights for each basic heading are available for all the countries.

A range of methods have been proposed in the literature to compute purchasing power parities for aggregation above the basic heading level. Some of the more popular ones are the Geary–Khamis (Geary, 1958; Khamis, 1970), Iklé (1972), country–product–dummy (CPD) (Rao, 1990, 2004, 2005; Diewert, 2005), Elteto–Koves–Szulc (EKS) (see, e.g. Rao, 2004) methods. In the recently completed 2005 round of the ICP, four aggregations methods have been used. At the basic heading level the CPD method was used in all regions except the OECD–EUROSTAT region where a modified EKS method was used. For aggregation at higher levels, the EKS method based on the Fisher binary index numbers was the method used in most regions. The Geary–Khamis method was recommended as an additively consistent aggregation method which was used in most regions except in the case of Africa where the Iklé method was used.

There has been a considerable amount of research on the properties of various aggregation methods. Balk (1996) compared the analytical properties of more than 10 different aggregation methods using the test approach. Hill (1997) provided a taxonomy of the aggregation methods. Diewert (1986) provided a framework for the comparison of various aggregation methods for international comparisons using the test approach. Despite the focus on the assessment of desirability of various index number formulae, research on the development of measures of reliability associated with PPPs computed using different formulae has been conspicuously absent.

Around the same time, when the developments in the area of international comparisons have been taking place, there has been a steady increase in research focusing on the stochastic approach to the construction of price and quantity index numbers with the primary aim of providing measures of reliability associated various index number formulae. Starting with an excellent evaluation of various approaches to index number construction by Ragnar Frisch (1936) and the consequent work of Theil (1967) on the stochastic approach to index numbers, the stochastic approach has been slowly developing. More recent interest on this approach started with the work of Clements and Izan (1981) and Selvanathan (1989), where the stochastic approach was used in the derivation of the Tornqvist index. Selvanathan and Rao (1992, 1994) demonstrated that the transitive multilateral index number proposed by Caves *et al*. (1982), which is essentially an EKS index constructed using binary Tornqvist indices, could be derived using the stochastic approach and the indices involved were shown to be generalized least squares estimators of parameters of an appropriately specified regression model. Rao and Selvanathan (1992) have shown that the stochastic approach could be

[1]See the *ICP Handbook* for more details (http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/ICPEXT/0,,contentMDK:20962711~menuPK:2666036~pagePK:60002244~piPK:62002388~theSitePK:270065,00.html).

used in deriving standard errors for PPPs computed using the Geary–Khamis method when the computations are conditioned on the knowledge of the international prices. Thus the approach of Rao and Selvanathan could be considered as a conditional–stochastic approach.

Along with these developments with respect to aggregation methods for the compilation of PPPs and the applications of the stochastic approach, there has been another strand of the stochastic approach gathering momentum. The CPD method, originally proposed by Summers (1973) as a tool for filling missing data in the price tableau used in international comparisons, has been found to play a significant role in the computation of PPPs at levels below and above the basic heading level. Kravis *et al*. (1982) recommend the use of the CPD method for the computation of PPPs below the basic heading level.[2] A modified version of the CPD method, referred to as the CPRD method, takes into account additional information concerning representativity of the product in the aggregation of price data below the basic heading level. However, it is only since the work of Rao (2004, 2005) that the CPD method has assumed a role in aggregation above the basic heading level. Rao (2005) has shown that the CPD method when applied along with expenditure share weight data (or the weighted CPD) results in PPPs that are identical to those derived using a method proposed in Rao (1990), which represents a modified version of the Geary–Khamis method. Following this line of research, Diewert (2005) has demonstrated that a number of commonly used methods can be derived using variants of the CPD method.

The main objective of this paper is to take the recent work of Rao and Diewert further and provide a stochastic framework for the derivation of a range of aggregation methods within the ICP. In particular, the paper focuses on four methods—the Geary–Khamis method, the Rao method, the Iklé method, and a variant of the Iklé method—and demonstrates that these methods can be shown to be estimators resulting from the use of the weighted maximum likelihood and the generalized method of moments procedures used in conjunction with the CPD method and a range of stochastic specifications for the disturbance term.[3]

The paper is organized as follows. Section 2 establishes the basic notation and provides an overview of the main aggregation methods considered in this paper. Section 3 briefly describes the CPD model used in international comparisons and shows how different systems are equivalent to the weighted maximum likelihood estimators of the parameters of the CPD model under different stochastic assumptions. Section 4 is devoted to a discussion on the method of deriving standard errors for the estimated PPPs. Section 5 focuses on the method of moments estimation of parameters of the CPD model. In Section 6 we present estimated PPPs and their standard errors using OECD international comparisons data for the 1996 benchmark year. The paper concludes with some remarks in Section 7.

[2]See Rao (2004) for more details on the CPD method and its properties.

[3]Derivation of the Fisher-based EKS method, which is the recommended aggregation method, and its variants is considered in detail in Rao (2009) where the main focus is on the generalized EKS and CPD methods.

2. Notation and Selected Multilateral Index Number Systems

Let p_{ij} and q_{ij} represent the price and the quantity of the j-th commodity in the i-th country, respectively, where $j = 1, \ldots, M$ indexes the countries and $i = 1, \ldots, N$ indexes the commodities. We assume that all the prices are strictly positive and all the quantities are non-negative, with the minimum condition that for each i, q_{ij} is strictly positive for at least one j; and for each j, q_{ij} is strictly positive for at least one i. Also, we let PPP_j denote purchasing power parity or the general price level in the j-th country relative to a numeraire country and P_i as the world average price for the i-th commodity. We also need the following systems of weights w_{ij} and w^*_{ij} in defining different systems of index numbers. These weights are defined as

$$(1) \qquad w_{ij} = \frac{p_{ij}q_{ij}}{\sum_{i=1}^{N} p_{ij}q_{ij}} \text{ and } w^*_{ij} = \frac{w_{ij}}{\sum_{j=1}^{M} w_{ij}}.$$

It is evident that $\sum_{i=1}^{N} w_{ij} = 1$ and $\sum_{j=1}^{M} w^*_{ij} = 1.$

The expenditure share weights, w_{ij}, reflect the relative importance of different commodities as measured by the budget shares. However, the second set of weights, w^*_{ij}, may best be described as the share of shares which reflect the importance of a given commodity i in country j relative to the importance attached to the commodity in all the countries involved in an international comparison. In contrast to w_{ij}, w^*_{ij}'s add up to unity over j (countries). The use of these weights arises naturally when defining cross-country or cross-regional average prices.[4]

We start with a description of the Geary–Khamis method, which is the first multilateral system to make use of the twin concepts of purchasing power parities (PPP_j) and international average prices (P_i).

Geary–Khamis Method

The Geary–Khamis multilateral system due to Geary (1958) and Khamis (1970)[5] is a popular method of aggregation for international comparisons as it provides additively consistent international comparisons. The Geary–Khamis system is defined by the following system of interdependent system of equations:

[4]Suppose we wish to find an average price, μ, which minimizes $\Sigma_j w_{ij} (p_{ij} - \mu)^2$. The objective here is to find a measure of central tendency of observed prices in different countries or regions which deviates the least from those commodities which are deemed to be important. As the average is over countries, the solution for this problem is a weighted average of observed prices in different countries with weights w^*_{ij}.

[5]Khamis has authored a number of papers that have delved deeply into various properties of the Geary–Khamis system.

$$(2)\quad PPP_j = \frac{\sum_{i=1}^{n} p_{ij}q_{ij}}{\sum_{i=1}^{n} P_i q_{ij}} \text{ for } j = 1, 2, \ldots, M; \text{and } P_i = \frac{\sum_{j=1}^{m}\left(p_{ij}q_{ij}/PPP_j\right)}{\sum_{j=1}^{m} q_{ij}} \text{ for } i = 1, 2, \ldots N.$$

For a given set of international prices, P_i, purchasing power parity of currency of country j, is defined as the ratio of value of the commodity bundle of country j evaluated, respectively, at the national prices, p_{ij}, and at the international prices, P_i. Similarly, for a given set of PPPs, international average prices are defined as the unit price derived from the total expenditure on commodity i across all countries and the total quantity of the commodity.

The simultaneous equation system in (2) has a solution that is unique up to a factor of proportionality. Given observed prices and quantity data from all the countries, the system is generally solved using an iterative procedure. Kravis *et al.* (1982) discuss various properties of the Geary–Khamis method; it remained as the principal aggregation method for international comparisons until the more recent phases of the ICP.[6] A major criticism of the method surrounds the definition of the international price, in (2), which is essentially a quantity weighted average of the observed prices in different countries. As a result the GK international prices tend to resemble those observed in richer countries and the real incomes of poorer countries tended to be overstated.[7]

We consider two aggregation methods which use the same framework as the Geary–Khamis (GK) method but are designed to address some of the main problems associated with the GK method.

Rao System for Multilateral Comparisons

Rao (1990) proposed a multilateral system derived through some modifications to the GK system. The Rao system replaces the quantity-share weights used in the definition of GK international prices by a system of weights that are based on expenditure shares. In addition, the system is defined using weighed geometric averages in the place of arithmetic averages used in the GK system. The system is defined as:

(3)

$$PPP_j = \prod_{i=1}^{N}\left(\frac{p_{ij}}{P_i} w_{ij}\right) \text{ for } j = 1, 2, \ldots M; \text{and } P_i = \prod_{j=1}^{M}\left(\frac{p_{ij}}{PPP_j} w_{ij}^{*}\right) \text{ for } i = 1, 2, \ldots N.$$

The system defined here is shown to have a non-trivial solution that is unique up to a factor of proportionality. In the case of binary comparisons, with $M = 2$, the

[6]The EKS method is now preferred as the principal aggregation method; the recently completed 2005 round of the ICP is based on the EKS method.

[7]This is usually referred to as the "Gerchenkron" effect.

Rao index is similar to the Tornqvist index.[8] The use of expenditure share weights reduced the likelihood of Gerchenkron effect present in the GK system. However, the Rao system is not additively consistent. We note here that both PPP_j's and P_i's are defined as a weighted geometric means of the price ratios, (p_{ij}/P_i), and national prices converted into a common currency unit, (p_{ij}/PPP_j), respectively.

Iklé System for Multilateral Comparisons

Iklé (1972) proposed an additively consistent system that is similar to the GK system which also makes use of the twin concepts of PPPs and international prices. Following Balk (1996), the Iklé (1972) system can be written in a form similar to equations (2) and (3). The system is given by:

(4)
$$\frac{1}{PPP_j} = \sum_{i=1}^{N}\left(\frac{P_i}{p_{ij}} w_{ij}\right) \text{ for } j = 1, 2, \ldots M; \text{and } \frac{1}{P_i} = \sum_{j=1}^{m}\left(\frac{PPP_j}{p_{ij}} w_{ij}^*\right) \text{ for } i = 1, 2, \ldots N.$$

The Iklé system also has a non-trivial solution which is unique up to a factor of proportionality. It is useful to note here that international prices, P_i, are defined as weighted harmonic means of prices observed in different countries after conversion to a common currency unit. Thus there is an element of commonality between the GK, Iklé, and Rao systems in that they use, respectively, weighted arithmetic, harmonic, and geometric averages of national prices. It should also be noted here that the first part of equation (4) defining *PPPj*'s is indeed identical to that used in defining *PPPj*'s in the Geary–Khamis system (2). These *PPPj*'s are essentially Paasche-type indices. The Iklé system has not been used in international comparisons until the 2005 ICP round.[9] It is useful to note here that equations defining the Iklé system in (4) are harmonic means instead of the geometric means used in defining the Rao-system.

A New Multilateral System with Expenditure Share Weighted Arithmetic Averages

Observing the statistical similarity between the Rao and Iklé systems, respectively, based on geometric and harmonic averages of the price relatives and prices, we consider a system defined using arithmetic averages defined using expenditure share weights used in the Rao and Iklé systems. The system is simply defined as:

(5)
$$PPP_j = \sum_{i=1}^{N}\left(\frac{p_{ij}}{P_i} w_{ij}\right) \text{ for } j = 1, 2, \ldots M; \text{and } P_i = \sum_{j=1}^{M}\left(\frac{p_{ij}}{PPP_j} w_{ij}^*\right) \text{ for } i = 1, 2, \ldots N.$$

The existence and uniqueness of solutions to system (5) is established in Hajargasht and Rao (2008). While the PPP_j's defined in the Rao system (3) are

[8]The binary index is essentially weighted geometric mean of price relatives where the weights are defined as harmonic means of expenditure share weights in the two countries.
[9]The Iklé method was used in the African region during the 2005 ICP round.

essentially Cobb–Douglas indices and those defined in the Iklé system are Paasche indices, no such simple interpretation can be accorded to the definition used for PPP_j's in equation (5). Therefore, there is no obvious bilateral index of the form in (5) used to define PPP_j. However, this may not be seen as a limitation as the international prices are not observed and no quantity and expenditure counterpart for these prices are known. As the current paper is focusing on the stochastic approach, it is sufficient to observe that the system defined in (5) has attractive stochastic properties similar to those associated with the weighted CPD and the Iklé system.

3. The Country–Product–Dummy Model and Multilateral Index Number Systems

So far we have described four systems of multilateral systems that are strongly linked in their conceptual framework with the Geary–Khamis system. In the next section we show that these systems can be derived as estimators of parameters of the CPD model under different distributional assumptions.

The CPD model was first proposed by Summers (1973) as a method of filling missing values in price data for international comparisons. It was also the preferred method of aggregation of price data below the basic heading level in international comparisons (Kravis *et al*., 1982). In the 2005 ICP round it has been the recommended method of aggregation below the basic heading level. The CPD model is gaining popularity as an aggregation method for aggregation above the basic heading level (see Rao, 2004, 2005; Diewert, 2005). The CPD model is now considered as the principal method of aggregation under the *stochastic approach*.

The CPD model postulates that the observed price of the *i*-th commodity in *j*-th country, p_{ij}, is the product of three components: the purchasing power parity (i.e. PPP_j); the price level of the *i*-th commodity relative to other commodities (i.e. P_i); and a random disturbance term u_{ij}. The CPD model is given by:

$$p_{ij} = P_i PPP_j u_{ij} \tag{6}$$

where u_{ij}'s are random disturbance terms which are independently and identically distributed. The parameters of the model (*PPP*s and *P*s) can be estimated from (6). The original model proposed by Summers (1973) simply transforms the model into a log-linear form and applies ordinary least squares to estimate the parameters. The estimated parameters are then used in filling any missing price observations. Rao (2005) showed that the Rao system defined in (3) is identical to the weighted least squares estimator of the parameters of the CPD model. This result has provided a useful link between the CPD model and aggregation methods above the basic heading level.

In this section we prove that the Rao and Iklé systems, and the new system can be derived as weighted maximum likelihood estimators of the parameters of the CPD model under different distributional assumptions for the disturbances, u_{ij}. We consider three different distributional specifications for the disturbances of the CPD model. In deciding on the specification, we take into consideration the fact that prices, hence the u_{ij}, are positive. Further, it is generally acknowledged that prices follow a skewed distribution. A natural choice of a distribution to represent the disturbances therefore would be the lognormal distribution. There is a considerable amount of literature surrounding the lognormal distribution, including the possibility of invoking the central limit theorem on observed prices if they are generated through multiplicative shocks. We also consider the gamma distribution, which is also commonly used in modeling positive random variables with a skewed distribution. The gamma distribution is more flexible than the lognormal distribution, which can represent a variety of distributions. We also use inverse-gamma distribution where the inverse of the price relatives follow a gamma distributions. We consider the issue of goodness-of-fit of these distributions in Section 6.

3.1. *CPD Model with Lognormal Disturbances and the Rao System*

We consider the case where u_{ij}'s are lognormally distributed. This means that $\ln u_{ij}$ is normally distributed, in this case with mean equal to zero and variance equal to σ^2. In this case we consider the CPD model in its log-linear form:

$$\ln p_{ij} = \ln P_i + \ln PPP_j + v_{ij} \text{ where } v_{ij} = \ln u_{ij} \sim N\left(0, \sigma^2\right).$$

This log-linear equation can be equivalently expressed in the form of a linear regression model:

$$\ln p_{ij} = \sum_{i=1}^{N} \eta_i D_i + \sum_{j=1}^{M} \pi_j D_j^* + v_{ij} \text{ where } \eta_i = \ln P_i \text{ and } \pi_j = \ln PPP_j \tag{7}$$

where D_i is the i-th commodity dummy variable which takes a value equal to 1 for commodity i and 0 otherwise; and D_j^* is the j-th country dummy variable which takes a value equal to 1 for a price observation belonging to country j and equal to 0 otherwise. Thus the explanatory variables in (7) are essentially country and product dummy variables and hence the model is known as the country–product–dummy model.

Under the lognormality of the disturbances, u_{ij}, in the original model, the maximum likelihood estimators of the parameters in the log-linear model are the same as the ordinary least squares estimators of the parameters since the disturbances, v_{ij}, are normally distributed. Now we consider the weighted regression model:[10]

$$\sqrt{w_{ij}} \ln p_{ij} = \sum_{i=1}^{N} \eta_i \sqrt{w_{ij}} D_i + \sum_{j=1}^{M} \pi_j \sqrt{w_{ij}} D_j^* + \sqrt{w_{ij}} v_{ij}. \tag{8}$$

Rao (2005) has shown that the least squares estimators of the parameters in the weighted CPD model (8) are identical to the solutions of the log-linear equations obtained from the Rao system in (3). Further, it can be easily shown that, under

[10]Applying OLS to equation (8) is equivalent to minimizing $\sum_i \sum_j w_{ij}\left(\ln p_{ij} - \sum_i \eta_i D_i - \sum_j \eta_j D_j^*\right)^2$ which is the weighted residual sum of squares.

lognormality of u_{ij} and normality of v_{ij} the weighted maximum likelihood estimators of the parameters in (7) are the same as the weighted least squares estimators obtained through (8).

The discussion here establishes the result that under the lognormality of the disturbances, the weighted maximum likelihood estimators of the parameters are identical to the PPP_j's and P_i's from the Rao (1990) system defined in (3).

3.2. *Gamma Distribution and the New Index*

Here we start with the CPD model and assume that u_{ij}'s follows a gamma distribution[11] as follows:

$$u_{ij} \sim Gamma(r, r) \tag{9}$$

where r is a parameter to be estimated. We combine the CPD model in (6) and the distributional assumption (9) to write:[12]

$$\frac{p_{ij}}{P_i PPP_j} \sim Gamma(r, r). \tag{10}$$

The choice of the *same* parameter r for the two parameters of the gamma distribution ensures that the expected value of the disturbance term is equal to 1.[13] Now we outline the weighted maximum likelihood method and establish the required equivalence.

Our purpose here is to estimate parameters (i.e. P_i, PPP_j, and r) using a maximum likelihood procedure. From the definition of the gamma density function we can easily show that

$$p_{ij} \sim \frac{r^r}{\Gamma(r)} \frac{p_{ij}^{\,r-1}}{P_i PPP_j^{\,r}} e^{-r\frac{p_{ij} r}{P_i PPP_j}}. \tag{11}$$

Therefore the log of density function can be written as:

$$LnL_{ij} \propto r \ln r - \ln \Gamma(r) + (r-1)\ln p_{ij} - \ln P_i - r \ln PPP_j - r\frac{p_{ij} r}{P_i PPP_j}. \tag{12}$$

We can proceed with this (log-)density function and obtain estimates of the parameters of interest using the standard maximum likelihood procedure, but we would like to incorporate the weights into the model as well. Use of weights is consistent with the standard index number approach of weighting price relatives by their

[11]The choice of the gamma distribution is guided by the fact that observed prices, after conversion to a common currency, have a skewed distribution. The assumption of lognormal distribution also implies a skewed distribution for prices.

[12]One may notice the close association of the proposed model to what is known as a generalized linear model with gamma distribution. A generalized linear gamma regression may be defined as $y_i/\mathbf{x}_i\boldsymbol{\beta} \sim Gamma(r,r)$ (see McCullagh and Nelder, 1989). Our model is a non-linear version of such a model.

[13]For further details on the lognormal, gamma, and inverse-gamma distributions used here, the reader is referred to Johnson *et al.* (1994).

expenditure shares. This is also the approach used by Rao (2005) where the weighted least squares method is employed.

One way of affecting this approach is to use a weighted likelihood estimation procedure. The weighted likelihood function is defined as

$$WL = \prod_{i=1}^{N}\prod_{j=1}^{M} L_{ij}^{w_{ij}/M} \tag{13}$$

and therefore the log-weighted likelihood function becomes

$$\ln WL = \sum_{i=1}^{N}\sum_{j=1}^{M} \frac{w_{ij}}{M} LnL_{ij}. \tag{14}$$

Then our weighted log-likelihood function becomes

$$\begin{aligned}\ln WL \propto (r-1)\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\ln p_{ij} - r\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\ln P_i - r\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\ln PPP_j - \\ r\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{p_{ij}w_{ij}}{P_i PPP_j} + r\ln r\left(\sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}\right) - \ln\Gamma(r)\sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}.\end{aligned} \tag{15}$$

Note that the above function may not represent a multivariate density function. Therefore, we do not interpret the estimation procedure as a maximum likelihood procedure. We rather interpret it as an M-estimation procedure (for more on M-estimators and their properties, see Wooldridge, 2002, chapter 12; Cameron and Trivedi, 2005, chapter 5).

Maximization of this objective function is not particularly difficult. The only potential problem is the presence of a gamma function in the likelihood function; however, most of the existing software, such as LIMDEP and GAUSS, can handle maximization of the functions containing gamma functions fairly easily.

The first order conditions from maximization of the above likelihood function are given by:

$$-\frac{r\sum_{j=1}^{M} w_{ij}}{P_i} + \frac{r}{P_i^2}\sum_{j=1}^{M}\frac{p_{ij}w_{ij}}{PPP_j} = 0 \qquad i = 1, \ldots, N$$

$$-\frac{r\sum_{i=1}^{N} w_{ij}}{PPP_j} + \frac{r}{PPP_j^2}\sum_{i=1}^{N}\frac{p_{ij}w_{ij}}{P_i} = 0 \qquad j = 1, \ldots, M$$

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\ln p_{ij} - \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\ln P_i - \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\ln PPP_j - \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{p_{ij}w_{ij}}{P_i PPP_j} + \\ M + M\ln r - M\frac{\partial}{\partial r}\ln\Gamma(r) = 0.\end{aligned}$$

There are $M+N+1$ first order conditions in as many unknown PPP_j ($j = 1,2, \ldots, M$), P_i ($i = 1,2, \ldots, N$), and r. After some algebraic manipulations, we can rewrite the above sets of equations as

(16)
$$P_i - \sum_{j=1}^{M} \frac{p_{ij} w_{ij}^*}{PPP_j} = 0$$

$$PPP_j - \sum_{i=1}^{N} \frac{p_{ij} w_{ij}}{P_i} = 0$$

$$\frac{\partial}{\partial r} \ln \Gamma(r) - \ln r = \frac{1}{M} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \ln p_{ij} - \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \ln P_i - \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \ln PPP_j \right).$$

The third equation simplifies to this form using the fact that taking sum over i in the second equation gives $\sum_{i=1}^{N} \sum_{j=1}^{M} \frac{p_{ij} w_{ij}}{P_i PPP_j} = M$. We observe that the first two equations in (16) are the same as the system of equations we introduced as the new system defined in (5); these equations do not depend upon the value of r.

Given Ps and PPPs (the existence and uniqueness of which have been proved elsewhere), it can be shown that there is a unique positive "r" which solves the third equation. To see this, note that $\frac{\partial}{\partial r} \ln \Gamma(r) - \ln r$ is an increasing and continuous function defined over positive values of "r" and its range changes from $-\infty$ to 0. Now if we show that the right-hand side of the third of (16) is always negative, the claim follows from a mean-value theorem. To prove the negativity, note that we can write

$$\ln \prod_{i=1}^{N} \prod_{j=1}^{M} \left[\frac{p_{ij}}{P_i PPP_j} \right]^{\frac{w_{ij}}{M}} < \ln \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{w_{ij}}{M} \frac{p_{ij}}{P_i PPP_j} = \ln(1) = 0.$$

The argument in the left is just the right-hand side of the third equation in (16). The inequality follows from geometric–arithmetic mean inequality.

Thus we have shown that the new multilateral system based on weighted arithmetic averages is identical to the weighted maximum likelihood estimator of the CPD model if the disturbances follow a gamma distribution with both the parameters set equal to each other.[14]

3.3. *Inverse-Gamma Distribution and the Iklé Index*

We follow the same approach as in Section 3.2 in the derivation of the Iklé index from the CPD model. In particular we show the weighted maximum likelihood estimator of the parameters of the CPD model when the disturbances follow

[14]It is possible to use differing parameter values in the gamma (r,r) distribution. However, in that case the expected value would be different from one.

inverse-gamma distribution. In order to use the inverse-gamma distribution, we rewrite the CPD model in (6) slightly differently. We use the reciprocal of the price and obtain:

$$\frac{1}{p_{ij}} = \frac{1}{P_i PPP_j} u_{ij} \tag{17}$$

where u_{ij}'s are random disturbance terms which are independently and identically distributed, and as before they are assumed to follow a gamma distribution:[15]

$$u_{ij} \sim Gamma(r, r) \tag{18}$$

where r is a parameter to be estimated. The model in equation (17) differs from the model in equation (10) mainly in the specification of the disturbance term and how it enters the equation. One of the possible advantages of this model is that we do not have the inverse relationship between variance of p_{ij} and w_{ij}. We combine (17) and (18) to write:

$$\frac{1}{p_{ij}} \propto \frac{r^r}{\Gamma(r)} \frac{\left(P_i PPP_j\right)^r}{p_{ij}^{\,r-1}} e^{-r\frac{P_i PPP_j}{p_{ij}}}. \tag{19}$$

Following the same procedure as we used in Section 3.2, we may obtain the likelihood function as

$$\begin{aligned} \ln L \propto & -(r-1)\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \ln p_{ij} + r\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \ln P_i + r\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \ln PPP_j - \\ & r\sum_{i=1}^{N}\sum_{j=1}^{M} \frac{P_i PPP_j w_{ij}}{p_{ij}} + r\ln r\left(\sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}\right) - \ln\Gamma(r)\sum_{i=1}^{N}\sum_{j=1}^{M} w_{ij}. \end{aligned} \tag{20}$$

Taking derivative with respect to **PPP** and **P** yields the Iklé system of equations

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

$$\frac{1}{P_i} = \sum_{j=1}^{M} \frac{PPP_j}{P_{ij}} w_{ij}^{*}.$$

Thus we have shown that the Iklé system is the same as the weighted likelihood estimators of the parameters of the CPD model under the assumption of inverse-gamma for the disturbances.

Results shown in Sections 3.1 to 3.3 establish that the Rao and Iklé systems, and the new system are all weighted maximum likelihood estimators of the

[15]Since the disturbance term in (17) is the reciprocal of the disturbance term in the original CPD model (6), the assumption in (18) is same as the assumption that disturbance term in (6) follows inverse-gamma distribution.

parameters of the CPD model that are distinguished by the differences in the distributions of the disturbance of the CPD model. Therefore, we have been able to show that all these index numbers belong to a class of index numbers based on the stochastic approach. Unfortunately, we have not been able to identify a distribution for the disturbance term under which the Geary–Khamis method could be derived. However, we show in Section 5 that the GK system can also be derived from the CPD model by showing that the GK system is equivalent to the method of moments estimator of the parameters of the CPD models. We will return to this shortly.

4. Computation of Standard Errors

We have emphasized that the advantage of the stochastic approach to index numbers and the use of CPD is to obtain standard errors for estimated indices. One might think that standard errors from conventional weighted least squares or weighted maximum likelihood provided by standard software can be used for this purpose. But such standard errors are not valid if these are not derived using proper expressions. Since we have shown that various systems of multilateral index numbers can be derived using the CPD model, it remains for us to derive the expressions to be used in deriving the standard errors. In order to derive standard errors for PPPs and international prices, P_i's, we make use of results available for M-estimators discussed in econometric literature.

We start with a general discussion of M-estimators and their variances. An M-Estimator $\hat{\boldsymbol{\theta}}$ is defined as an estimator that maximizes an objective function of the following form (see, e.g. Cameron and Trivedi, 2005):

$$Q_N(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} h_i(y_i, \mathbf{x}_i; \boldsymbol{\theta}) \tag{22}$$

where y_i and $\mathbf{x}_i$ represent dependent and independent variables, respectively. $\boldsymbol{\theta}$ is the vector of parameters to be estimated. The function Q is the same as the weighted likelihood function in logarithmic form given in equations (15) and (20).

Following Cameron and Trivedi (2005), it has been shown that $\hat{\boldsymbol{\theta}}$ has the following asymptotic distribution:

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow{d} \aleph[0, \mathbf{A}_0^{-1}\mathbf{B}_0\mathbf{A}_0^{-1}]$$

where

$$\mathbf{A}_0 = \text{plim}\frac{1}{N}\sum_{i=1}^{N}\left.\frac{\partial^2 h_i}{\partial\boldsymbol{\theta}'\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0} \tag{23}$$

$$\mathbf{B}_0 = \text{p}\lim\frac{1}{N}\sum_{i=1}^{N}\left.\frac{\partial h_i}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_0}\left.\frac{\partial h_i}{\partial\boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}_0}.$$

In practice, a consistent estimator can be obtained as:

$$\mathbf{VAR}(\hat{\boldsymbol{\theta}}) = \frac{1}{N}\hat{\mathbf{A}}^{-1}\hat{\mathbf{B}}\hat{\mathbf{A}}^{-1} \tag{24}$$

where

$$\hat{\mathbf{A}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial^2 h}{\partial\boldsymbol{\theta}'\partial\boldsymbol{\theta}}\bigg|_{\hat{\boldsymbol{\theta}}} \tag{25}$$

$$\hat{\mathbf{B}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial h_i}{\partial\boldsymbol{\theta}}\bigg|_{\hat{\boldsymbol{\theta}}}\frac{\partial h_i}{\partial\boldsymbol{\theta}'}\bigg|_{\hat{\boldsymbol{\theta}}}. \tag{26}$$

In some special cases like the maximum likelihood or non-linear least squares with homoscedastic errors, it can be shown that $\mathbf{A}_0^{-1} = -\mathbf{B}_0$. In such cases the variance formula can be simplified to

$$\mathbf{VAR}(\hat{\boldsymbol{\theta}}) = -\frac{1}{N}\hat{\mathbf{A}}^{-1}. \tag{27}$$

Many software programs use this formula as their default standard error formula. But in case of the problem studied in this paper, this formula leads to incorrect standard errors for the estimated parameters and we must use the more general formula given by (23).

For example, if we apply formula (27) to the estimates from a weighted least squares regression, we obtain following formula:

$$\mathbf{VAR}(\hat{\boldsymbol{\theta}}) = \hat{\sigma}^2(\mathbf{X}'\boldsymbol{\Omega}\mathbf{X})^{-1} \tag{28}$$

where $\boldsymbol{\Omega}$ is a diagonal matrix with weights on its diagonal which coincide with the standard formula for weighted least squares when there is heteroscedasticity in error term. However, the correct formula for the variance estimator to be used in the case where we used weighted least squares when the disturbances are homoscedastic, is given by:

$$\mathbf{VAR}(\boldsymbol{\theta}) = \hat{\sigma}^2(\mathbf{X}'\boldsymbol{\Omega}\mathbf{X})^{-1}(\mathbf{X}'\boldsymbol{\Omega}'\boldsymbol{\Omega}\mathbf{X})(\mathbf{X}'\boldsymbol{\Omega}\mathbf{X})^{-1} \tag{29}$$

where $\hat{\sigma}^2$ is obtained from the unweighted regression. This formula is similar to that suggested in Rao (2004) for the computation of standard errors for the weighted CPD method.

The results presented here are quite useful in conducting statistical inference on the unknown PPPs. The estimated *PPPs* are asymptotically normally distributed. Therefore, it would be possible to obtain confidence intervals and to conduct hypothesis tests using standard normal tables. It would be useful if simple algebraic expressions of the standard errors associated with *PPPs* are derived.[16] The expression in (29) implies that the magnitude of the standard errors would depend upon the estimate of variance σ^2. From the specification of the CPD model, large values of σ^2 imply large deviations from the *law of one price* implicit in the CPD model.

In Section 6 we present estimated PPPs based on price data from the OECD and using different methods along with their standard errors derived under different stochastic assumptions discussed in Section 3. Before that we turn to the derivation of the GK system from the CPD model.

5. Derivation of the Geary–Khamis System Using the CPD Model

We recall that the Geary–Khamis system in equation (2) is given by:

$$PPP_j = \frac{\sum_{i=1}^{N} p_{ij}q_{ij}}{\sum_{i=1}^{N} P_i q_{ij}} \text{ for } j = 1,2, \ldots, M; \text{ and } P_i = \frac{\sum_{j=1}^{M}\left(p_{ij}q_{ij}/PPP_j\right)}{\sum_{j=1}^{M} q_{ij}} \text{ for } i = 1,2, \ldots N.$$

In the past there have been several attempts to cast the GK method in a stochastic framework so that standard errors can be derived.

One of the early attempts was due to Rao and Selvanathan (1992) who used a stochastic specification where *PPPj*'s are identified as parameters of a regression model where the international prices P_i's are assumed to be *known*. Thus the standard errors are *conditional* on the full knowledge of the international prices. But in practice, international prices are also unknown and determined simultaneously with the unknown *PPP*s. Recently, Diewert (2005) derived the Geary–Khamis bilateral index using the stochastic approach based on the CPD method for the case of binary comparisons. The Diewert approach also makes use of several steps which makes it difficult to derive the standard errors for the *PPP*s.

In this paper, we show that the Geary–Khamis *PPP*s and the international prices, P_i's, in the multilateral case are the method of moments estimators of the parameters of the CPD model in equation (6) discussed in Section 3.1. In particular, the approach used here recognizes the non-additive nature of the CPD model and proposes the method of moments approach. In Section 5.1 we discuss how a non-additive non-linear system of equations can be estimated using the method of moments. Section 5.2 applies this approach to the CPD model, which is a non-additive model, and shows how the arithmetic and the Geary–Khamis indices can be derived using this approach. A numerical illustration which presents the GK *PPP*s and their standard errors is included in Section 6.

[16]Derivation of closed form expressions for the standard errors is fairly involved and beyond the scope of this paper.

5.1. *Estimation of Non-Additive Non-Linear Models*

In establishing a relationship between the GK method and the CPD model, we consider the CPD model as a non-additive model and then look at the problem of estimation of the parameters of the non-additive model using the method of moments estimation technique.

Consider the following non-linear regression model:

$$r(y_i, \mathbf{x}_i, \boldsymbol{\beta}) = u_i \tag{30}$$

where y_i represents the dependent variable, u_i represents the random errors, $r(y_i, \mathbf{x}_i, \boldsymbol{\beta})$ is a non-linear function, $\mathbf{x}_i$ is a $1 \times L$ vector, $\boldsymbol{\beta}$ is a $K \times 1$ column vector, and $i = 1, \ldots, N$ indexes the number of observations; we also assume that $E(u_i) = 0$. We make a further assumption that the model is non-additive,[17] which means it cannot be written as

$$y_i - g(\mathbf{x}_i, \boldsymbol{\beta}) = u_i. \tag{31}$$

Parameters of an additive model can be estimated using a non-linear least squares approach, but it can be shown that the use of least square criterion does not provide consistent estimators for non-additive models (see, e.g. Cameron and Trivedi, 2005).

How can a non-additive model be estimated? We consider the method of moments estimation of the parameters of the model. An obvious starting point is to base the estimation of parameters in (31) on the moment conditions $E(\mathbf{X}'\mathbf{u}) = \mathbf{0}$, where $\mathbf{X}$ is the $N \times L$ matrix containing $\mathbf{x}_i$'s and $\mathbf{u}$ is an $N \times 1$ vector containing u_i's. However, other moment conditions can be used. More generally we can base the estimation on the following K moment conditions:

$$E\left(\mathbf{R}(\mathbf{x}, \boldsymbol{\beta})'\mathbf{u}\right) = \mathbf{0} \tag{32}$$

where $\boldsymbol{R}$ is an $N \times K$ 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 following sample moment conditions:

$$\frac{1}{N}\mathbf{R}\left(\mathbf{X}, \hat{\boldsymbol{\beta}}\right)'\mathbf{r}\left(\mathbf{y}, \mathbf{X}, \hat{\boldsymbol{\beta}}\right) = \mathbf{0}. \tag{33}$$

This estimator is asymptotically normal with variance matrix

$$Var\left(\hat{\boldsymbol{\beta}}_{MM}\right) = \hat{\sigma}^2\left[\hat{\mathbf{D}}'\hat{\mathbf{R}}\right]^{-1}\hat{\mathbf{R}}'\hat{\mathbf{R}}\left[\hat{\mathbf{R}}'\hat{\mathbf{D}}\right]^{-1} \tag{34}$$

where $\hat{\mathbf{D}} = \left.\dfrac{\partial \mathbf{r}(\mathbf{y}, \mathbf{X}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}'}\right|_{\hat{\boldsymbol{\beta}}}$, $\hat{\mathbf{R}} = \mathbf{R}\left(\mathbf{X}, \hat{\boldsymbol{\beta}}\right)$ and $\hat{\sigma}^2 = \dfrac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{N}$.

[17]It is easy to check that the CPD model is a non-additive model using the definition.

The main issue in the above estimation problem is the specification of the moment conditions defined by $\mathbf{R}(\mathbf{X},\boldsymbol{\beta})$. It has been shown (see, e.g. Davidson and MacKinnon, 2004) that the most efficient choice for the moment conditions is

$$\text{(35)} \qquad \mathbf{R}(\mathbf{X},\boldsymbol{\beta})^* = E\left[\frac{\partial \mathbf{r}(\mathbf{y},\mathbf{X},\boldsymbol{\beta})'}{\partial \boldsymbol{\beta}}\middle|\mathbf{X}\right].$$

In general the expectation term in the right-hand side cannot be derived unless we make very strong distributional assumptions, but fortunately for the type of models we consider in this paper it is tractable.

5.2. *Estimation of PPPs Under the Optimal Choice of Moment Conditions and Standard Errors Using MOM*

To obtain PPPs and their standard errors based on an the CPD model using MOM, we follow Rao (2005) and Diewert (2005) again to postulate that the observed price of the *j*-th commodity in the *i*-th country, p_{ij}, is the product of three components: the purchasing power parity (i.e. PPP_j); the price level of the *j*-th commodity relative to other commodities (i.e. P_i); and a random disturbance term as follows:

$$\text{(36)} \qquad p_{ij} = P_i PPP_j u_{ij}^*$$

where $u_{ij}*$s are random disturbance terms which are independently and identically distributed.[18] We also assume that $E(u_{ij}^*) = 1$. The model in equation (36) can be written in the following equivalent form:

$$\text{(37)} \qquad \frac{p_{ij}}{P_i PPP_j} - 1 = u_{ij}$$

with $E(u_{ij}) = 0$. This is now in the form of a non-additive non-linear regression model as introduced in the previous section, and therefore we can use the estimation method in the previous section. Using the theory discussed in the previous section, the equations to be solved can be written as

$$\text{(38)} \qquad \frac{1}{nm}\mathbf{R}'\mathbf{r} = \mathbf{0}$$

where $\mathbf{R}'$ is an $(n + m) \times (n \times m)$ matrix; it can be shown that the most efficient choice of R according to (35) is defined as follows:

[18]We use u_{ij}^* instead of u_{ij} in order to facilitate the specification of the non-additive model shown in (37).

$$
\mathbf{R}' = \mathbf{E}\begin{bmatrix}
\begin{matrix} -\frac{p_{11}}{P_1^2 PPP_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{p_{n1}}{P_n^2 PPP_1} \end{matrix} &
\begin{matrix} -\frac{p_{12}}{P_1^2 PPP_2} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{p_{n2}}{P_n^2 PPP_2} \end{matrix} &
\cdots &
\begin{matrix} -\frac{p_{1m}}{P_1^2 PPP_m} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{p_{nm}}{P_n^2 PPP_m} \end{matrix} \\
\begin{matrix} -\frac{p_{11}}{P_1 PPP_1^2} & \cdots & -\frac{p_{n1}}{P_n PPP_1^2} \\ 0 & & \\ 0 & & \end{matrix} &
\begin{matrix} 0 & \cdots & \\ & \ddots & \\ & & 0 \end{matrix} & &
\begin{matrix} & & 0 \\ & & 0 \\ -\frac{p_{1m}}{P_1 PPP_{1m}^2} & \cdots & -\frac{p_{nm}}{P_n PPP_m^2} \end{matrix}
\end{bmatrix}
$$

and

(39) $$r_{ij} = \frac{p_{ij}}{P_i PPP_j} - 1$$

considering the fact that

(40) $$E\left[\frac{p_{ij}}{P_i PPP_j}\right] = 1.$$

We can write the equations in the following matrix form:

$$
\begin{bmatrix}
\begin{matrix} -\frac{1}{P_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{1}{P_n} \end{matrix} &
\begin{matrix} -\frac{1}{P_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{1}{P_n} \end{matrix} &
\cdots &
\begin{matrix} -\frac{1}{P_n} & & \bigcirc \\ & \ddots & \\ \bigcirc & & -\frac{1}{P_n} \end{matrix} \\
\begin{matrix} -\frac{1}{PPP_1} & \cdots & -\frac{1}{PPP_1} \\ 0 & & \\ 0 & & \end{matrix} &
\begin{matrix} 0 & \cdots & \\ & \ddots & \\ & & 0 \end{matrix} & &
\begin{matrix} & & 0 \\ & & 0 \\ -\frac{1}{PPP_m} & \cdots & -\frac{1}{PPP_m} \end{matrix}
\end{bmatrix}
\begin{bmatrix}
\frac{p_{11}}{P_1\, PPP_1} - 1 \\
\frac{p_{12}}{P_2\, PPP_1} - 1 \\
\vdots \\
\vdots \\
\vdots \\
\frac{p_{nm}}{P_n\, PPP_m} - 1
\end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix}.
$$

We can write the normal equations as follows:

$$\text{(41)}\qquad \begin{cases} -\dfrac{1}{P_i}\displaystyle\sum_{j=1}^{M}\left(\dfrac{p_{ij}}{P_i PPP_j}-1\right)=0 \\ -\dfrac{1}{PPP_j}\displaystyle\sum_{i=1}^{N}\left(\dfrac{p_{ij}}{P_i PPP_j}-1\right)=0 \end{cases} \Rightarrow \begin{cases} P_i=\dfrac{1}{m}\displaystyle\sum_{j=1}^{m}\left(\dfrac{p_{ij}}{P_i PPP_j}\right) \\ PPP_j=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\left(\dfrac{p_{ij}}{P_i PPP_j}\right) \end{cases}.$$

According to the theory in the previous section, the variance for the estimated price indexes can be obtained by

$$\text{(42)}\qquad Var\left(\hat{\boldsymbol{\beta}}_{MM}\right)=\hat{\sigma}^2\left[\hat{\mathbf{D}}'\hat{\mathbf{R}}\right]^{-1}\hat{\mathbf{R}}'\hat{\mathbf{R}}\left[\hat{\mathbf{R}}'\hat{\mathbf{D}}\right]^{-1}$$

where

$$\mathbf{D}'=\begin{bmatrix} -\frac{p_{11}}{P_1^2PPP_1} & 0 & -\frac{p_{12}}{P_1^2PPP_2} & 0 & \cdots & -\frac{p_{1m}}{P_1^2PPP_m} & 0 \\ 0 & \ddots & 0 & \ddots & \cdot & 0 & \ddots \\ & -\frac{p_{n1}}{P_n^2PPP_1} & & -\frac{p_{n2}}{P_n^2PPP_2} & \cdots & & -\frac{p_{nm}}{P_n^2PPP_m} \\ -\frac{p_{11}}{P_1\,PPP_1^2} \;\cdot\cdot\; -\frac{p_{n1}}{P_n\,PPP_1^2} & & 0 & & & & 0 \\ 0 & & & \ddots & & & 0 \\ 0 & & & & 0 & -\frac{p_{1m}}{P_1\,PPP_{1m}^2} \;\cdots & -\frac{p_{nm}}{P_n\,PPP_m^2} \end{bmatrix}.$$

So far we have not introduced weights in our price index. One way of doing this is to define the **R** matrix as follows:

$$\mathbf{R}'=\begin{bmatrix} -\frac{w_{11}}{P_1} & 0 & -\frac{w_{12}}{P_1} & 0 & \cdots & -\frac{w_{1m}}{P_n} & 0 \\ 0 & \ddots & 0 & \ddots & \cdot & 0 & \ddots \\ & -\frac{w_{n1}}{P_n} & & -\frac{w_{n2}}{P_n} & \cdots & & -\frac{w_{nm}}{P_n} \\ -\frac{w_{11}}{PPP_1} \;\cdots\; -\frac{w_{n1}}{PPP_1} & & 0 & & & & 0 \\ 0 & & & \ddots & & & 0 \\ 0 & & & 0 & & -\frac{w_{1m}}{PPP_m} \;\cdots & -\frac{w_{nm}}{PPP_m} \end{bmatrix}.$$

This definition for the **R** matrix results in the following system of equations which coincides with the weighted version of the arithmetic index:

$$\left\{\begin{aligned} PPP_j &= \sum_{i=1}^{N}\left(\frac{p_{ij}}{P_i}w_{ij}\right) \\ P_i &= \sum_{j=1}^{M}\frac{p_{ij}}{PPP_j}w_{ij}^{*} \end{aligned}\right. . \tag{43}$$

This set of equations is the same equations that defined the new system based on the expenditure share weighted arithmetic means to define PPPs and P_i's. This is exactly the arithmetic index introduced earlier in equation (5) in Section 2.

5.3. *Derivation of the Geary–Khamis PPPs and Standard Errors*

Consider again estimation of the following non-additive CPD model:

$$\frac{p_{ij}}{P_i PPP_j} - 1 = u_{ij}.$$

As we discussed in the previous sections, we can base our estimation on the following moment conditions:

$$E[\mathbf{R}'\mathbf{u}] = \mathbf{0}$$

and accordingly following sample moment conditions

$$\frac{1}{nm}\mathbf{R}'\mathbf{r} = \mathbf{0}.$$

Different definitions for **R** can lead to different estimators. As long as **R** is not correlated with **u** the estimator is consistent. We make a slight modification in the definition of **R** in the previous section as follows:

$$\mathbf{R}' = \begin{bmatrix} \begin{matrix} -\frac{1}{P_1} & & O \\ & \ddots & \\ O & & -\frac{1}{P_n} \end{matrix} & \begin{matrix} -\frac{1}{P_1} & & O \\ & \ddots & \\ O & & -\frac{1}{P_n} \end{matrix} & \cdots & \begin{matrix} -\frac{1}{P_1} & & O \\ & \ddots & \\ O & & -\frac{1}{P_n} \end{matrix} \\ \begin{matrix} -\frac{P_1}{PPP_1} & \cdots & -\frac{P_n}{PPP_1} \end{matrix} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ 0 & \cdots & 0 & \begin{matrix} -\frac{P_1}{PPP_m} & \cdots & -\frac{P_n}{PPP_m} \end{matrix} \end{bmatrix}.$$

It is easy to see that **R** is not correlated with **u** because **P** and **PPP** are constant parameters of the model to be estimated. (Note also that P_i's are close to one and therefore this matrix does not differ very much from the one in the last section.) This definition for **R** results in the following equations:

$$\begin{cases} \sum_{i=1}^{n} P_i\left(\dfrac{p_{ij}}{P_i PPP_j}-1\right)=0 \\ \sum_{j=1}^{N} 1\left(\dfrac{p_{ij}}{P_i PPP_j}-1\right)=0 \end{cases} \Rightarrow \begin{cases} PPP_j = \dfrac{\sum_{i=1}^{n} p_{ij}}{\sum_{i=1}^{n} P_i} \\ P_i = \dfrac{1}{m}\sum_{j=1}^{m}\left(\dfrac{p_{ij}}{PPP_j}\right) \end{cases}.$$

But this is the unweighted Geary–Khamis price index. We can derive the quantity-weighted price index by defining

$$\mathbf{R}' = \begin{bmatrix} \begin{matrix} -\frac{q_{11}}{P_1} & & 0 \\ & \ddots & \\ 0 & & -\frac{q_{n1}}{P_n} \end{matrix} & \begin{matrix} -\frac{q_{12}}{P_1} & & 0 \\ & \ddots & \\ 0 & & -\frac{q_{n2}}{P_n} \end{matrix} & \cdots & \begin{matrix} -\frac{q_{1m}}{P_n} & & 0 \\ & \ddots & \\ 0 & & -\frac{q_{nm}}{P_n} \end{matrix} \\ \begin{matrix} -\frac{q_{11}P_1}{PPP_1} & \cdots & -\frac{q_{n1}P_n}{PPP_1} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{matrix} & \begin{matrix} 0 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & 0 \end{matrix} & & \begin{matrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ -\frac{q_{1m}P_1}{PPP_m} & \cdots & -\frac{q_{nm}P_n}{PPP_m} \end{matrix} \end{bmatrix}.$$

This results in the following system of equations:

$$\begin{cases} PPP_j = \dfrac{\sum_{i=1}^{n} p_{ij}q_{ij}}{\sum_{i=1}^{n} P_i q_{ij}} \\ P_i = \dfrac{\sum_{j=1}^{m}\left(p_{ij}q_{ij}/PPP_j\right)}{\sum_{j=1}^{m} q_{ij}} \end{cases}$$

which is identical to the equations that define the Geary–Khamis system given in equation (2) in Section 2. Thus it is clear that the GK PPPs and P_i's are the method of moments (weighted) estimators of the parameters of the CPD model.

As usual the standard errors for the estimated indexes can be obtained using the following formula:

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

where $\mathbf{D}_{ij}$'s are the same as in the previous section.

The result established in this section provides for the very first time a proper derivation of the GK system using a stochastic approach. The MOM estimator derived here relates to the estimation of both PPPs and Ps simultaneously. This is more general than the partial approach used in Rao and Selvanathan (1992). This result also provides a method of estimating standard errors for PPPs from the GK method.

6. Empirical Application Using OECD Data

In this section we present estimated PPPs and their standard errors derived using the three methods of aggregation discussed in the paper and the 1996 OECD data. The price information that we have is in the form of PPPs at the basic heading level for 158 basic headings, with the US dollar used as the numeraire currency. In addition we have expenditure, in national currency units, for each basic heading in all the OECD countries. These nominal expenditures provide the expenditure share data used in deriving the weighted maximum likelihood estimators under alternative stochastic specification of the disturbances.

For weighted CPD estimates we have used the weighted least squares methodology as explained in Rao (2005). For Iklé and the new index we used the weighted maximum likelihood approach described in Section 2.

Results shown in Table 1 clearly demonstrate the feasibility and comparability of the new approaches to the estimation of PPPs. As can be seen, PPPs and their standard errors based on CPD, Iklé, and the new index are all numerically close to each other. An additional phenomenon to note is that the PPPs based on the weighted CPD (or from the lognormal specification for the disturbances) appear to be bounded by PPP estimates from the new index and the Iklé index. However, this is only a coincidence and when a different country (e.g. Australia) is used as the reference country, no special patterns emerged.

The PPPs reported here, from different methods, are very similar to each other. As the countries selected are developed OECD countries, it is generally acknowledged that *PPPs* are usually close to the market exchange rates. The relative precision of the estimated *PPPs* can be compared by looking at the ratio of the standard error to the estimated *PPP*. These ratios are generally higher for lower income countries within the OECD, such as Turkey and Portugal, and countries with relatively higher price levels, such as Switzerland and Norway. However, further analysis and work involving countries at various stages of development is necessary before any concrete conclusions can be drawn. Such work along with the derivation of analytical expressions for the standard errors is a topic for future research.

TABLE 1
MLE Estimates of PPPs and SEs

	MLE Estimates					
	New Index		CPD		Iklé	
Country	PPP	SE	PPP	SE	PPP	SE
GER	1.887	0.136	2.034	0.144	2.187	0.147
FRA	6.092	0.429	6.554	0.455	7.035	0.466
ITA	1,425.96	109.727	1,504.02	115.509	1,584.381	119.196
NLD	1.921	0.150	2.056	0.155	2.205	0.156
BEL	35.491	2.577	37.890	2.698	40.450	2.728
LUX	33.578	2.488	35.816	2.618	38.191	2.700
UK	0.603	0.043	0.642	0.044	0.682	0.045
IRE	0.637	0.051	0.669	0.055	0.696	0.060
DNK	8.525	0.586	9.131	0.615	9.762	0.631
GRC	180.470	13.452	188.482	13.891	196.640	14.005
SPA	112.414	8.304	118.546	8.606	124.799	8.738
PRT	126.043	10.400	129.037	10.994	130.317	12.002
AUT	12.770	0.881	13.730	0.928	14.728	0.948
SUI	2.050	0.168	2.183	0.177	2.320	0.180
SWE	9.424	0.686	10.075	0.720	10.758	0.742
FIN	6.159	0.432	6.598	0.453	7.070	0.462
ICE	86.828	7.000	89.541	6.975	92.329	6.810
NOR	8.807	0.684	9.238	0.736	9.642	0.764
TUR	6,304.23	579.128	6,321.42	544.907	6,357.003	506.991
AUS	1.264	0.099	1.333	0.103	1.407	0.104
NZL	1.464	0.111	1.530	0.113	1.596	0.115
JAP	182.031	13.622	187.429	14.282	192.392	14.780
CAN	1.168	0.090	1.229	0.094	1.295	0.096
USA	1.0		1.0		1.0	

Table 2 shows the estimated PPPs and their standard errors based on: (i) the arithmetic index using MOM; and (ii) the Geary–Khamis method derived as an MOM estimator. The standard errors of the arithmetic index based on the maximum likelihood estimation (MLE) approach discussed in Sections 4 and 5 of this paper are also presented.

The results presented in the table are consistent with prior expectations. The standard errors for the arithmetic index using the generalized method of moments (GMM) show that this is slightly more efficient than MLE. This could be because GMM is robust to the choice of distribution for the error term; the standard errors for the Geary–Khamis using the method proposed here are higher than the other two, which is expected because it is not the most efficient estimator based on our stochastic specification.

Which Disturbance Specification?

It is clear from the empirical results presented here that it is possible to derive PPPs from different methods by simply varying the distribution of the disturbance term. Or alternatively, use a method of moments estimator which does not rely on any distributional assumptions.

It is useful to review the differences between different distribution specifications. The disturbances are positive and skewed random variables. While

TABLE 2
Estimates of PPPs and SEs

	Arithmetic Index	MOM SE Arithmetic	MLE SE Arithmetic	GK Index	MOM SE GK
GER	1.887	0.109442	0.136	2.08316	0.15474
FRA	6.092	0.606755	0.429	6.679491	0.516194
ITA	1,425.96	79.25337	109.727	1,537.168	129.5046
NLD	1.921	0.11156	0.150	2.032161	0.156602
BEL	35.491	1.946125	2.577	38.70436	2.700867
LUX	33.578	2.454269	2.488	36.7877	3.446165
UK	0.603	0.036311	0.043	0.679564	0.053761
IRE	0.637	0.037709	0.051	0.657754	0.056569
DNK	8.525	0.591807	0.586	9.457703	0.872669
GRC	180.470	9.271153	13.452	187.3352	13.14857
SPA	112.414	7.726502	8.304	122.1712	10.59001
PRT	126.043	6.56711	10.400	124.7745	9.307088
AUT	12.770	0.731266	0.881	14.40264	1.098328
SUI	2.050	0.146331	0.168	2.220059	0.179608
SWE	9.424	0.726701	0.686	10.56069	1.024583
FIN	6.159	0.404593	0.432	6.895726	0.638499
ICE	86.828	6.142211	7.000	90.02853	9.473389
NOR	8.807	0.457666	0.684	9.119335	0.764748
TUR	6,304.23	393.9744	579.128	5,967.556	549.1221
AUS	1.264	0.08598	0.099	1.351173	0.106996
NZL	1.464	0.106893	0.111	1.545069	0.140098
JAP	182.031	12.52263	13.622	179.0048	15.83708
CAN	1.168	0.085695	0.090	1.271441	0.115112
USA	1.0			1	

lognormal is a standard specification used in economics applications, use of gamma and other distributions is now commonplace within the class of generalized linear models. We make the following points regarding these distributions. For large values of r, the gamma and lognormal distributions are pretty close and they converge to each other as r gets larger. Therefore, for many practical problems with large values of r, it may not matter which one to choose, but for smaller values of "r" they can be different. For such values of "r," gamma tends to give most of the mass to smaller values of the stochastic variable.

For a sample drawn from a gamma distribution, the maximum likelihood estimator for the mean is the "arithmetic mean," and for sample drawn from a lognormal distribution it is the geometric mean. As expected, for most values of "r" these two become pretty close, but for smaller values of "r" the two averages can be very different.[19]

We have not yet established a formal test procedure which can be used in selecting a distribution from lognormal, gamma, and inverse-gamma distributions based on the observed price data. In order to provide an intuitive explanation as to how we may choose between these distributions, we make a comparison of the residuals from the CPD model in (2) with different specifications. We simply run

[19]We conducted a small simulation experiment where we generated a sample of 1000 from a gamma distribution with $r = 1$ and calculated the arithmetic mean which was close to one (as expected), but the geometric mean was close to 0.5. In this case the use of lognormal could severely bias the estimates.

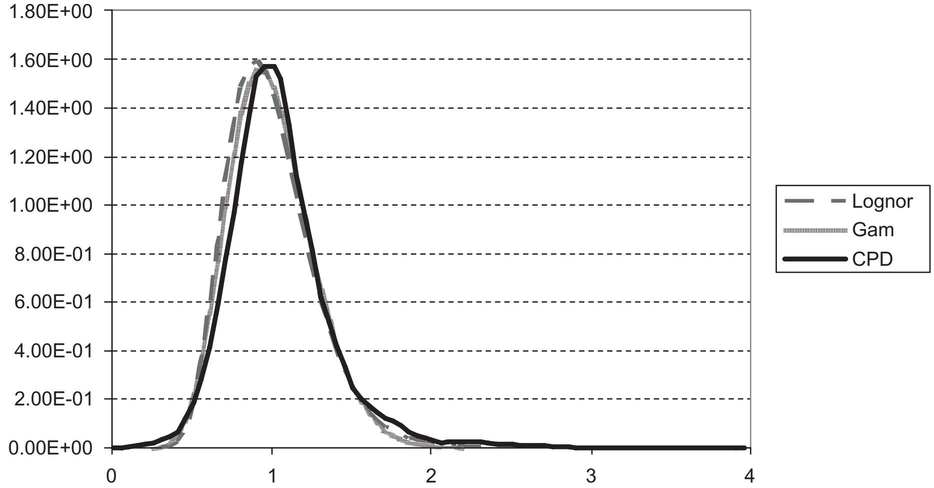

Figure 1. Distribution of the Disturbances of the CPD Model

the CPD model as a log-linear model in prices and compute the residuals. These residuals are presented in Figure 1, labeled as CPD. On the same graph, we also plot the disturbances drawn from a lognormal distribution and from a gamma distribution (r,r).[20]

The density function under the CPD model simply represents the residuals derived using the OLS estimators of the parameters of the CPD model without any distributional assumptions. The distributions implied by lognormal and gamma distributions are also presented. From the figure it appears that the gamma distribution provides a better approximation to the disturbances from the OLS. An implication of this is that if we were to select the gamma distribution to represent the distribution of the disturbances of the CPD model, then we should be using the arithmetic version of the GK system using expenditure share weights. However, this is an issue that requires further research.

7. Concluding Remarks

The paper has proposed a straightforward extension to two known multilateral methods due to Iklé (1972) and Rao (1990). The new index uses weighted arithmetic averages to define PPPs and international prices, *Pi*'s, instead of harmonic and geometric averages used, respectively, in the Iklé and Rao specifications. The paper has also established that all three indexes can be shown to be the weighted maximum likelihood estimators of the CPD model when the disturbances follow lognormal, gamma, or inverse-gamma distributions, respectively. Derivation of the indices using the stochastic approach makes it possible to derive appropriate standard errors for the Iklé index and the new index proposed here. Further, given that all these indexes are generated by the same CPD model, but

[20]In plotting the gamma distribution we make use of the estimated value of r equal to 12.

with alternative disturbance specifications, it allows us to test for the distributional assumptions underlying these three methods and use such specification tests to choose between alternative methods. Further work is necessary to see whether it is possible to explore other specifications for the distribution of the disturbance and the index number formulae resulting from such specifications. The paper also outlines the approach necessary to compute the true standard errors of PPPs when weighted maximum likelihood methods are used.

The paper has also shown that the commonly used Geary–Khamis PPPs can be derived from the CPD model and the stochastic approach described here. In particular, the GK PPPs are shown to be weighted method of moments (MOM) estimators of the parameters of the CPD model.

References

Balk, Bert M., "A Comparison of Ten Methods for Multilateral International Price and Volume Comparison," *Journal of Official Statistics*, 12, 199–222, 1996.

Cameron, C. and P. Trivedi, *Microeconometrics: Methods and Applications*, Cambridge University Press, New York, 2005.

Caves, D. W., L. R. Christensen, and W. E. Diewert, "Multilateral Comparisons of Output, Input and Productivity using Superlative Index Numbers," *Economic Journal*, 92, 73–86, 1982.

Clements, K. W. and H. Y. Izan, "A Note on Estimating Divisia Index Numbers," *International Economic Review*, 22, 745–7, 1981.

Davidson, R. and J. G. MacKinnon, *Econometric Theory and Method*, Oxford University Press, New York, 2004.

Diewert, W. E., "Microeconomic Approaches to the Theory of International Comparisons," NBER Technical Working Paper No. 53, 1986.

———, "Weighted Country Product Dummy Variable Regressions and Index Number Formulae," *Review of Income and Wealth*, 51, 561–70, 2005.

Frisch, R., "Annual Survey of General Economic Theory: The Problem of Index Numbers," *Econometrica*, 4, 1–38, 1936.

Geary, R. C., "A Note on the Comparison of Exchange Rates and Purchasing Power between Countries," *Journal of the Royal Statistical Society, Series A (General)*, 121, 97–9, 1958.

Hajargasht, G. and D. S. Prasada Rao, "Systems of Index Numbers using the Stochastic Approach," Paper presented at the 2008 World Congress on National Accounts and Economic Performance Measures for Nations, Key Bridge Marriott Hotel, Arlington, Virginia, May 12–17, 2008.

Hill, R. J., "A Taxonomy of Multilateral Methods for Making International Comparisons of Prices and Quantities," *Review of Income and Wealth*, 43, 49–69, 1997.

Johnson, N. L., S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions, Vol. I*, 2nd edition, John Wiley & Sons, 1994.

Iklé, D. M., "A New Approach to the Index Number Problem," *Quarterly Journal of Economics*, 86, 188–211, 1972.

Khamis, S. H., "Properties and Conditions for the Existence of a New Type of Index Numbers," *Sankhya, Series B*, 32, 1970.

Kravis, I. B., A. W. Heston, and R. Summers, *World Product and Income: International Comparisons of Real Gross Domestic product*, Johns Hopkins University Press, Baltimore, 1982.

McCullagh, P. and J. A. Nelder, *Generalized Linear Models*, Chapman and Hall, London, 1989.

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

———, "The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power Parities in the ICP," Paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, Vancouver, June 30–July 3, 2004.

———, "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, 2005.

———, "Generalised Elteto-Koves-Szulc (EKS) and Country-Product-Dummy (CPD) methods for International Comparisons," in D. S. Prasada Rao (ed.), *Purchasing Power Parities: Recent Advances in Methods and Applications*, Edward Elgar, Aldershot, 86–120, 2009.

Rao, D. S. Prasada and E. A. Selvanathan, "Computation of Standard Errors for Geary–Khamis Parities and International Prices: A Stochastic Approach," *Journal of Business and Economic Statistics*, 10, 101–15, 1992.

Selvanathan, E. A., "A Note on the Stochastic Approach to Index Numbers," *Journal of Business and Economic Statistics*, 7, 471–4, 1989.

Selvanathan, E. A. and D. S. Prasada Rao, "An Econometric Approach to the Construction of Generalised Theil–Tornqvist Indices for Multilateral Comparisons," *Journal of Econometrics*, 54, 335–46, 1992.

———, *Index Numbers: A Stochastic Approach*, The University of Michigan Press, Ann Arbor, 1994.

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

Theil, H., *Economics and Information Theory*, North-Holland, Amsterdam, 1967.

Wooldridge, J., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA, 2002.