International Journal of Engineering and Applied Sciences (IJEAS)
ISSN: 2394-3661, Volume-3, Issue-4, April 2016

Review: The Stochastic Approach and Systems of Index Numbers

El-Hoseeny Abd El-bar Rady, **Aly Ibrahim Mohamed**, **Ahmed Amin El-sheikh,**
Hanan Mohamed Mostafa Mohamed

Abstract— **The main objective of the paper is to review a number of widely used multilateral index numbers for international comparisons of purchasing power parities (PPPs) and real incomes that can be derived using the stochastic approach. The paper discuss 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 paper also presents the method of moments (MOM) as an approach to estimate PPPs under the stochastic approach and shows how the Geary–Khamis system of multilateral index numbers is a method of moments estimator of the parameters of the CPD model.**

Index Terms— **Index Numbers - Stochastic Approach - multilateral index numbers - method of moments – country product dummy (CPD) model**

I. Introduction

The stochastic approach, used in derived the purchasing power parities (ppps) which are essentially spatial price index numbers that provide measures of price level differences across countries or regions within a country.
The Country Product Dummy (CPD) model represented a simple regression approach to measures price level differences in different countries or regions within a country.
The CPD model proposed by Summers (1973), as a model for filling missing price data in the context of international comparisons. Subsequently the CPD model used in the first few phases of the International Comparison Project (ICP) mainly as a tool for aggregating price data under the basic heading level.
Rao (1995), proposed Weighted- CPD as a link between the CPD model and the geometric variant of the Geary-Khamis model. In addition, it considered a generalization of the CPD model, which used for aggregation above the basic heading level.

$$\ln p_{ij} = \pi_1 D_1 + \pi_2 D_2 + \ldots + \pi_M D_M + \eta_1 D_1^* + \eta_2 D_2^* + \ldots + \eta_N D_N^* + u_{ij} \qquad (1)$$

Where:

p_{ij} is price of commodity i in the country j.

π_j is the $\ln ppp_j$, and $j = 1,2,\ldots,M$.

η_i is the $\ln p_i$, and $i = 1,2,\ldots,N$.

$$D_j = \begin{cases} 1 & \text{if price } p_{ij} \text{ observation belongs to country j} \\ 0 & \text{o/w} \end{cases}$$

$$D_i^* = \begin{cases} 1 & \text{if price } p_{ij} \text{ observation refers to } i^{-th} \text{ commodity} \\ 0 & \text{o/w} \end{cases}$$

M and N is number of countries and commodities respectively.

u_{ij}'s are random disturbance terms which are independently and identically distributed.

II. The Country Product Dummy (CPD) Method:

Rao (2004), used the (CPD) model to represent multiple regression approach to measure price level differences in different countries. The CPD model postulates that the p_{ij}, is the product of three components:

$$p_{ij} = p_i \cdot ppp_j \cdot V_{ij} \qquad (2)$$

Where:

p_i is the commodity i^{-th} price, ppp_j is the purchasing power parity is the price level of the i^{-th} commodity relative to other commodities price is p_i and V_{ij}'s are a random disturbance terms which are iid distributed. The parameters of the model (ppp_j and p_i) can be estimated:

$$\ln p_{ij} = \ln p_i + \ln ppp_j + \ln V_{ij} \qquad (3)$$

$$\ln p_{ij} = \eta_i + \pi_j + u_{ij} \qquad (4)$$

Where, η_i and π_j are defined as before and u_{ij} are

El-Hoseeny Abd El-bar Rady, Dep. of Applied Statistics andEconometrics, Institute of Statistical Studies and Research Cairo University**, Egypt**
Aly Ibrahim Mohamed,Deputy of Agricultural Economics Research Institute Agricultural Economics, Research Institute Agricultural Research Center**, Egypt**
Ahmed Amin El-sheikh, Dep. of Applied Statistics and Econometrics, Institute of Statistical Studies and Research Cairo University**, Egypt.**
Hanan Mohamed Mostafa Mohamed, Deputy of Agricultural Economics Research Institute,Agricultural Economics, Research Institute Agricultural Research Center, Egypt

random disturbance which are iid normally distributed $u_{ij} : N(0, \sigma^2)$.

It is possible to express P_i relative to a reference country (say country 1), then π_j represents the purchasing power parity ppp_j of country j which showing the number of country j currency units that have the same purchasing power as one unit of currency of country 1. The ppp for country, j is given by:

$$ppp_j = \exp\left[\pi_j\right] \tag{5}$$

This system solved by imposing a linear restriction on the unknown parameters. For $\alpha_1 = 0$ is the restriction imposed, it can be easily shown that, for each j=2,…, M

$$ppp_j = \exp(\pi_j) = \prod_{i=1}^{N}\left[\frac{p_{ij}}{p_i}\right]^{1/N} \tag{6}$$

Using Eq(6),to make a comparisons of price levels between two countries c and d, represented by ppp_{cd} can be derived as:

$$PPP_{cd} = \frac{PPP_d}{PPP_c} = \prod_{i=1}^{N}\left[\frac{p_{id}}{p_{ic}}\right]^{1/N} \tag{7}$$

The price level comparison, PPP_{cd}, in Eq(7) satisfied the transitivity property and Eq (7) is equivalent to the EKS (Elteto-Koves-Szulc) index number.

III. The CPD Regression Model

The standard CPD formulation is multiple regression model that regresses logarithm of observed prices on a set of dummy variables representing the commodity and country to which a given price observation refers to as in the Eq (1) this model can be written as:

$$y_{ij} = x_{ij}\beta + u_{ij} \tag{8}$$

Where:

$$x_{ij} = \left[D_1, D_2, \ldots, D_M, D_1^*, D_2^*, \ldots, D_N^*\right]$$

$$\beta = \left[\pi_1, \pi_2, \ldots, \pi_M, \eta_1, \eta_2, \ldots, \eta_N\right]'$$

The values of the dummy variables are determined at the observation ij of NM observations (for j =1,2,…M and i =1,2,…N).

The model can be rewritten as:

$$y - X\beta = u \tag{9}$$

Eq (9) is a general regression model with NM observations and (N+M) explanatory variables. The matrix X has rank (N+M-1) as the first M dummy variables and the last N dummy variables sum to the same vector with elements equal to 1.

The restriction $\pi_1 = 0$ imposed in Eq (8), this restriction implies that the currency of country 1 taken as the numerator and PPPs of all other countries are expressed relative to the currency of country 1. The model under this restriction is essentially the same as in Eq (9) except that the first column of X and the first element of β are dropped from the model. The restricted model is then given by:

$$y - X^*\beta^* = v \tag{10}$$

Under the assumption of independently and iid random disturbances, the best linear unbiased estimator of β^* and the associated covariance matrix is given by:

$$\hat{\beta}^* = \left(X^{*\prime}X^*\right)^{-1} X^{*\prime} y \tag{11}$$

$$\mathrm{Var}\left(\hat{\beta}^*\right) = \sigma^2\left(X^{*\prime}X^*\right)^{-1} \tag{12}$$

Using the special structure of the matrix X*, consisting of various dummy variables as columns, it can be shown that:

$$X^{*\prime}X^* = \begin{bmatrix} N\, I_{(M-1)\times(M-1)} & J_{(M-1)\times N} \\ J_{N\times(M-1)} & M\, I_{N\times N} \end{bmatrix} \tag{13}$$

Where I is an identity matrix and J is a rectangular matrix of appropriate order whose elements are all equal to 1. Using the special structure of $X^{*\prime}X^*$:

$$\left(X^{*\prime}X^*\right)^{-1} = \begin{bmatrix} \frac{1}{N} I_{(M-1)\times(M-1)} + \frac{1}{N} J_{(M-1)\times(M-1)} & \frac{1}{NM-(N+M)} J_{(M-1)\times N} \\ \frac{1}{NM-(N+M)} J_{N\times(M-1)} & \frac{1}{M} I_N + \frac{M-1}{MN} J_{N\times N} \end{bmatrix} \tag{14}$$

Given this expression, it can be shown that:

$$\hat{\pi}_j = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{p_{ij}}{p_i}\right] \tag{15}$$

Rao (2004), also proved that $\hat{\pi}_j$ is an unbiased estimator of π_j and its variance is given by:

$$V\left(\hat{\pi}_j\right) = \frac{2}{N}\hat{\sigma}^2 \tag{16}$$

To estimate the variances of these parameters, the least squares residuals defined for each observation y_{ij} as:

$$v_{ij} = y_{ij} - \hat{\pi}_j - \hat{\gamma}_i \tag{17}$$

Since, v_{ij}'s are assumed to be iid with mean zero and variance σ^2.

An unbiased estimator of σ^2 is given by:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} e_{ij}^2}{MN-(M+N-1)} \quad (18)$$

It is possible to derive the estimated variance of each element of $\hat{\beta}$. Eq (18) is illustrated that the estimated standard errors of the $\hat{\pi}_j$'s are the same for all j=1,2,…M. Therefore, there is little that distinguishes the estimators of the logarithmic price level differences in terms of standard errors. This is mainly due to the fact that all commodities are priced in all the countries and disturbances in equation, v_{ij} are i id.

IV. Systems of Price Index Numbers for International Comparisons Based on Stochastic Approach:

These systems are a class of index numbers for international price comparisons, for example Rao system, Geary-Ghamis system (GK), Hajargasht-Rao, Elteto-Koves-Szulc (EKS), and IKLE system. Diewert (2005) has demonstrated that a number of commonly used formulae can be derived using the CPD model. Although, Rao (2005) established that the Rao method for computing PPPs is equivalent to the weighted CPD model.

Rao, (2005) supposed that p_{ij} and q_{ij} are the price and the quantity of the i-th commodity in the j-th country respectively, where j =1,2,…,M indexes the countries and i =1,2,…,N indexes the commodities. Also defined ppp_j as purchasing power parity or the general price level in the j-th country relative to a numerator country , P_i is the world average price for the i-th commodity and supposed that the following system of weights W_{ij} and W_{ij}^{*} is defining different systems of index numbers. These weights are defined as:

$$W_{ij} = \frac{p_{ij}q_{ij}}{\sum_{i=1}^{N} p_{ij}q_{ij}}, \qquad \sum_{i=1}^{N} W_{ij} = 1, \quad (19)$$

$$W_{ij}^{*} = \frac{W_{ij}}{\sum_{j=1}^{M} W_{ij}}, \qquad \sum_{j=1}^{M} W_{ij}^{*} = 1, \quad (20)$$

Lemma(1):

The Rao system is the weighed geometric averages in the place of arithmetic averages used in the GK system as follow:

$$p_i = \prod_{j=1}^{M}\left(\frac{p_{ij}}{ppp_j}\right)^{W_{ij}^{*}} \qquad i=1,2,\ldots,N \quad (21)$$

$$ppp_j = \prod_{j=1}^{N}\left(\frac{p_{ij}}{p_i}\right)^{W_{ij}} \qquad j=1,2,\ldots M \quad (22)$$

These two equations of the Rao-system are equivalent to the weighted least squares estimate for the weighted CPD model.

V. Hajargasht and Rao Multilateral System of Index Number:

Hajargasht and Rao (2010), proposed their index number system , (H-R) with expenditure share weighted, it is arithmetic averages. The system is:

$$p_i = \left(\sum_{j=1}^{M}\frac{p_{ij}}{ppp_j}w_{ij}^{*}\right) \qquad i=1,2,\ldots,N \quad (23)$$

$$ppp_j = \left(\sum_{i=1}^{N}\frac{p_{ij}}{p_i}w_{ij}\right) \qquad j=1,2,\ldots,M \quad (24)$$

Where, p_i , ppp_j , p_{ij} , W_{ij} , W_{ij}^{*} and D_i , D_j^{*} are defined as before.

Lemma(2): If the disturbances terms u_{ij} distributed gamma (r,r) in the derivation of Hajargasht and Rao multilateral system with expenditure share weighted arithmetic averages, then the H-R multilateral system is equivalent to the weighted maximum likelihood estimators of parameters of the CPD model.

VI. IKLE Multilateral System of Index Number:

The IKLE (1972) system, is an additive consistent system that is similar to the GK system and makes use of the two concepts of ppp_j and international prices p_i :

$$\frac{1}{ppp_j} = \left(\sum_{i=1}^{N}\frac{p_i}{p_{ij}}w_{ij}\right) \qquad j=1,2,\ldots,M \quad (25)$$

$$\frac{1}{p_i} = \left(\sum_{j=1}^{M}\frac{ppp_j}{p_{ij}}w_{ij}^{*}\right) \qquad i=1,2,\ldots,N \quad (26)$$

Where: p_i , ppp_j , p_{ij} , W_{ij} , W_{ij}^{*} and D_i , D_j^{*} are defined as before.

Lemma(3): Hajargasht and Rao (2010), proved that if the u_{ij} disturbance term distributed gamma (r,r) the IKLE system of index number is the MLE of the parameters from the CPD model take the form:

$$\frac{1}{p_{ij}} = \frac{1}{P_i ppp_j} \cdot u_{ij} \qquad (27)$$

Where u_{ij} are random disturbance terms which are iid and as before they are assumed to follow a gamma distribution.

VII. The Geary-Khamis Multilateral System of Index Number:

The Geary-Khamis (Geary, 1958; Khamis, 1970) multilateral system, is a system of index number which based on the same two concepts ppp_j of currencies, and averages of prices - P_i . Computations based on a simultaneous equation system as follow:

$$ppp_j = \frac{\sum_{i=1}^{N} p_{ij} q_{ij}}{\sum_{i=1}^{N} P_i q_{ij}} \qquad (28)$$

$$P_i = \frac{\sum_{j=1}^{M} \left(p_{ij} q_{ij} / ppp_j \right)}{\sum_{j=1}^{M} q_{ij}} \qquad (29)$$

Where: P_i , ppp_j , p_{ij} and q_{ij} defined as before

Hajargasht and Rao (2010), proved that the G-K multilateral system of index number is equivalent to the method of moments (MOM) estimator of the parameters of the CPD models as follow.

$$p_{ij} = P_i \; ppp_j \; u^*_{ij} \qquad (30)$$

Where u^*_{ij} are random disturbance terms, which are iid, also assumed that Eq (30) can be written in the following equivalent form.

$$\frac{p_{ij}}{P_i \; ppp_j} - 1 = u_{ij} \qquad (31)$$

With $E(u_{ij}) = 1$ the form of Eq (31) is a non-additive nonlinear regression model.

A. *Estimation of non-additive nonlinear models:*

Hajargasht and Rao (2010) considered that the CPD model is a non-additive model and then used the method of moment estimation technique to estimate the parameters. Assume the nonlinear regression model:

$$r(y_i, x_i, \beta) = u_i \qquad , i=1,2,\ldots,N \qquad (32)$$

Where:

y_i is the dependent variable, u_i is the random errors, x_i is a L × 1 vector and β is a 1 × K column vector, $r(y_i, x_i, \beta)$ is a nonlinear function and i =1,...., N indexes the number of observations and also assumed that $E(u_i) = 0$. The starting point to base the estimation of parameters in on the moment conditions is :

$$E(X'u) = 0 \qquad (33)$$

Where, X is N×L matrix containing x_i's and u is an N×1 vector containing u_i's.

Generally the estimation based on the following K moment conditions:

$$E(R(x,\beta)'u) = 0 \qquad (34)$$

Where: R is a N× K vector of functions of X and β .

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:

$$\frac{1}{N}(R(x, \hat\beta)' r(y, x, \hat\beta)) = 0 \qquad (35)$$

This estimator is asymptotically normal with variance matrix:

$$var(\hat\beta_{MM}) = \hat\sigma^2 \left[\hat D'\hat R\right]^{-1} \hat R'\hat R \left[\hat R'\hat D\right]^{-1} \qquad (36)$$

Where:

$$\hat D = \frac{\partial r(y, x, \beta)}{\partial \beta}\Big|_{\beta=\hat\beta} \quad , \hat R = R(X, \hat\beta) \quad , \hat\sigma^2 = \frac{\hat u'\hat u}{N} \qquad (37)$$

B. *Estimation of PPPs under the optimal choice of moment conditions and standard errors using MOM:*

To postulate the observed price of j-th commodity in i-th country, p_{ij} is the product of three components, the purchasing power parity ppp ,the price level of the j-th commodity relative to other commodities, P_i and a random disturbance term as in Eq (30) as:

$$p_{ij} = P_i \; ppp_j \; u^*_{ij}$$

Where: u^*_{ij} are random disturbance terms, which are iid.

Rewrite this model in the following equivalent form as in Eq (31) as:

$$\frac{p_{ij}}{P_i \; ppp_j} - 1 = u_{ij}$$

With $E(u_{ij}) = 1$ this equation is a non-additive nonlinear regression model. This model to be solved can be written as:

$$\frac{1}{NM} R' r = 0 \tag{38}$$

Where, R is an $(M+N)(M+N)$ matrix and it can be shown that most efficient choice of R is defined as:

$$R' = E\begin{bmatrix} -\frac{p_{11}}{p_1^2 ppp_1} & & -\frac{p_{12}}{p_1^2 ppp_2} & \cdots & -\frac{p_{1M}}{p_1^2 ppp_M} & \\ & -\frac{p_{N1}}{p_N^2 ppp_1} & & -\frac{p_{N2}}{p_N^2 ppp_2} & & -\frac{p_{NM}}{p_N^2 ppp_M} \\ -\frac{p_{11}}{p_1^2 ppp_1} \cdots & -\frac{p_{1M}}{p_1\, ppp_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{p_{21}}{p_1\, ppp_1} \cdots & -\frac{p_{2M}}{p_1\, ppp_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{p_{1M}}{p_1\, ppp_M} \cdots & -\frac{p_{NM}}{p_N\, ppp_M} \end{bmatrix} \tag{39}$$

Since,

$$E\left[\frac{p_{ij}}{p_i\, ppp_j}\right] = 1 \text{ and } r_{ij} = \frac{p_{ij}}{p_i\, ppp_j} - 1\text{, then:}$$

$$R'r = \begin{bmatrix} -\frac{1}{p_1} & 0 & -\frac{1}{p_1} & 0 & -\frac{1}{p_N} & 0 \\ 0 & -\frac{1}{p_N} & 0 & -\frac{1}{p_N} & 0 & -\frac{1}{p_N} \\ -\frac{1}{ppp_1} \cdots & -\frac{1}{ppp_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{ppp_2} \cdots & -\frac{1}{ppp_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{ppp_M} \cdots & -\frac{1}{ppp_M} \end{bmatrix} \begin{bmatrix} \frac{p_{11}}{p_1\, ppp_1} - 1 \\ \frac{p_{12}}{p_2\, ppp_1} - 1 \\ \vdots \\ \vdots \\ \frac{p_{NM}}{p_N\, ppp_M} - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \tag{40}$$

Then the system of normal equations has the following forms:

$$\begin{cases} p_i = \frac{1}{M}\sum_{j=1}^{M}\left[\frac{p_{ij}}{p_i\, ppp_j}\right] \\ ppp_j = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{p_{ij}}{p_i\, ppp_j}\right] \end{cases} \tag{41}$$

To impose weights in the price index, the R matrix will be defined as:

$$R' = \begin{bmatrix} -\frac{w_{11}}{p_1} & 0 & -\frac{w_{12}}{p_1} & 0 & -\frac{w_{1M}}{p_1} & 0 \\ 0 & -\frac{w_{N1}}{p_N} & 0 & -\frac{w_{N2}}{p_N} & & -\frac{w_{NM}}{p_N} \\ -\frac{w_{11}}{ppp_1} \ldots\ldots & -\frac{w_{N1}}{ppp_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{w_{12}}{ppp_1} \ldots\ldots & -\frac{w_{N2}}{ppp_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{w_{1M}}{ppp_M} \ldots\ldots & -\frac{w_{NM}}{ppp_M} \end{bmatrix} \tag{42}$$

Then the system of normal equations has the following forms:

$$\begin{cases} p_i = \sum_{j=1}^{M}\left[\frac{p_{ij}}{p\,p\,p_j} w^*_{ij}\right] \\ ppp_j = \sum_{i=1}^{N}\left[\frac{p_{ij}}{p_i} w_{ij}\right] \end{cases} \tag{43}$$

This set of equations is the same as equations that defined the new system based on the expenditure share weighted arithmetic means to define ppp_j's and p_i's.

C. Estimation of PPPs in Gk system under the optimal choice of moment conditions and standard errors using MOM in:

To derive weighted Geary-Khamis (GK) price index, the quantity weighted price index R can be defined as:

$$R' = \begin{bmatrix} -\frac{q_{11}}{p_1} & & -\frac{q_{12}}{p_1} & & -\frac{q_{1M}}{p_N} & \\ & -\frac{q_{N1}}{p_N} & & -\frac{q_{N2}}{p_N} & & -\frac{q_{NM}}{p_N} \\ -\frac{q_{11}p_1}{ppp_1}\ldots\ldots & -\frac{q_{N1}p_1}{ppp_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{q_{12}p_1}{ppp_2}\ldots\ldots & -\frac{q_{N2}p_N}{ppp_2} & 0 & 0 \\ & & & & \ddots & \\ 0 & 0 & 0 & 0 & -\frac{q_{1M}p_1}{ppp_M}\ldots\ldots & -\frac{q_{NM}p_N}{ppp_M} \end{bmatrix} \tag{44}$$

Then the system of normal equations has the following forms:

$$\begin{cases} p_i = \sum\limits_{j=1}^{M} \left[\dfrac{p_{ij} q_{ij} \,/\, ppp_j}{\sum\limits_{j=1}^{M} q_{ij}} \right] \\ ppp_j = \sum\limits_{i=1}^{N} \left[\dfrac{p_{ij} q_{ij}}{\sum\limits_{j=1}^{M} q_{ij}} \right] \end{cases} \qquad (45)$$

Which is identical to the equations of the Geary-Khamis system given in Eq (28) and Eq(29). Thus it is clear that the G-K PPPs and P_i's are the method of moments (weighted) estimators of the parameters of the CPD model.

REFERENCES

[1] Diewert, W. E. "Weighted Country Product Dummy Variable Regressions and Index Number Formulae," *Review of Income and Wealth*,2005,51, 561–70.

[2] Hajargasht, G. and Rao, D.S.Prasada, "Systems of Index Numbers for International Price Comparisons Based on the Stochastic Approach", *Review of Income and Wealth*, 2010, 56(1), S32–S58.

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

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

[5] 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)*,1958, 121, 97–9.

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

[7] "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, 2004.

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

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