International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

A MULTIVARIATE WEIGHTING APPROACH FOR CONSUMER PRICE INDEXING

Ayimah, J.C.
Ho Technical University;
Mettle, F.O.
Nortey, E.N.N.
University of Ghana, Legon, Accra;
Minkah, R.
University of Ghana,

ABSTRACT: *This study proposed a multivariate weighting system, which allows for the computation of both base and current year weights from price datasets conveniently. The proposed multivariate weighting system adopts a factor analysis methodology aimed at mitigating the widely reported weight and formula biases associated with CPI formulation, with an advantage of retaining variability in each price variable in the final index. The precisions of the expenditure-based weight (current weighting system) and that of the multivariate weighting system were assessed. It was found that the proposed weighting system performed better in terms of the proportion of indices falling within the ideal range. This result was consistent for 1000 bootstrap samples generated from 10 different multivariate data samples simulated for this purpose. The paper recommends the multivariate weighting system, because it is easy to implement and cost effective, as an alternative for the expenditure-based weighting system noted in literature for biases.*

KEYWORDS: index numbers; consumer price index; multivariate statistics; classical factor analysis

INTRODUCTION

Consumer Price Index (CPI) measures changes in prices of goods and services and it is used in the determination of inflation and living conditions in a country. The CPI is an index number primarily used to measure changes in prices of some set of goods consumed by households in a quest to satisfy their needs and wants (International Labour Office & Turvey, 2004). In the 1980s, the well-known Laspeyres' and Paasche's indices were proposed. In the composition of these indices, data on prices, quantity purchases, and expenditure incurred by consumers are required. Then from the quantities and expenditures, weights attributable to items are determined. To capture the dynamics of price changes over a period, a superlative index (in which weights capture the dynamics of price changes over a period) is the ideal. However, this is often difficult to compute in real-time due to the fact that data on current quantity and expenditure by consumers are not readily available at the time of index formulation. This challenge, results in the popular weight and formula biases associated with CPI computations worldwide.

This paper seeks to propose a multivariate measure that can be used to formulate an alternative weighting system using solely price data, so that, the variability explained in each price variable would be retained in the final index. Essentially, the utilisation of this weighting system curtails the existence of the popular weight and formula biases associated with CPI formulation.

Subsequent sections of the paper are organised under the following headings: Background of the Study, Methodology, Results, Discussions, Conclusions and Recommendations.

Theoretical Underpinning

Four popular index formulas are known from literature: The "modified" Laspeyres' index, whose weights are obtained from base year expenditure data:

$$CPI_L = \frac{\sum_{i=1}^{p} w_{ib}\left(\frac{P_{ic}}{P_{ib}}\right)}{\sum_{i=1}^{p} w_{ib}}; \qquad (1)$$

the "modified" Paasche's index, with weights computed from current year expenditure data:

$$CPI_P = \frac{\sum_{i=1}^{p} w_{ic}\left(\frac{P_{ic}}{P_{ib}}\right)}{\sum_{i=1}^{p} w_{ic}}; \qquad (2)$$

then the Fisher's Ideal superlative index, a geometric mean of the Laspeyres' and the Paasche's indices:

$$CPI_F = \left(\frac{\sum_{i=1}^{p} w_{ib}\left(\frac{P_{ic}}{P_{ib}}\right)}{\sum_{i=1}^{p} w_{ib}}\right)^{1/2} \times \left(\frac{\sum_{i=1}^{p} w_{ic}\left(\frac{P_{ic}}{P_{ib}}\right)}{\sum_{i=1}^{p} w_{ic}}\right)^{\frac{1}{2}} \qquad (3)$$

and the Drobish-Bowley's index, another superlative index, which is an arithmetic mean of the Laspeyres' and the Paasche's indices:

$$CPI_{DB} = \frac{\left(\frac{\sum_{i=1}^{p} w_{ib}\left(\frac{P_{ic}}{P_{ib}}\right)}{\sum_{i=1}^{p} w_{ib}}\right) + \left(\frac{\sum_{i=1}^{p} w_{ic}\left(\frac{P_{ic}}{P_{ib}}\right)}{\sum_{i=1}^{p} w_{ic}}\right)}{2}; \qquad (4)$$

where

P_{iC} is the price of the i^{th} item in the current year,
P_{iB} is the price of the i^{th} item in the base year,
w_{iC} weight of the i^{th} item in the current year,
w_{iB} weight of the i^{th}item in the base year.
p number of items in the basket.

The debate on the biases in CPI was mostly due to biases in the formulation of weights for use in the above formulas (Boskin et al., 1996; Malik et al., 2014). The current weighting system is expenditure-based, computed once in a base year as:

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

$$w_{ib} = \frac{E_{ib}}{\sum_{i=1}^{p} E_{ib}}, \qquad (5)$$

Or as:

$$w_{ib} = \left(\frac{E_{ib}}{\sum_{i=1}^{p} E_{ib}}\right)100. \qquad (6)$$

where E_{ib} is the expenditure made by households on the i^{th} item.

The concern over the decades of debate has been the fact that, E_{ib} is a product of household expenditure survey (HES) for which reason w_{ib} is computed once in several years. Simply, w_{ib} may not be representative over a period. Again, the huge cost involved in conducting HES makes it even more prohibitive to carry out HES yearly (International Labour Office, & Turvey, 2004; Srivastava & Srivastava, 2003). Essentially, the assumption that consumer taste and standard of living remains the same over a period cannot be realistic in certain price environments (Malik et al., 2014; Milana, 2009). To this end, the popular weight and formula biases created in the formulation of CPI is the root-cause of the CPI problem (Malik et al., 2014) and for which the Laspeyres' price index is deemed unrepresentative (Khalid & Asghar, 2010). The ideal situation is to have superlative indices (Afriat, 2005, 1977; Hicks, 1956; Konus, 1924; Samuelson & Swamy, 1974; Samuelson, 1974; Swamy, 1984), which are composed of and bounded by the Laspeyres' and Paasche's indices. Clearly, the issue about the unrepresentativeness of the current CPI weighting system (Bryan & Cecchetti, 1993; Cecchetti, 1996; Clark, 1999; Reed & Rippy, 2012) had gained much attention in literature.

In view of these challenges, CPI is computed in the USA for urban dwellers only (McCully et al., 2007) and the personal consumption expenditures price index (PCEPI) is computed and reported as well (Khalid & Asghar, 2010; McCully et al., 2007). The PCEPI computes weights from sales data taken from shops, supermarkets, malls, etc. PCEPI weighting system suggests that the system can only apply when data on sales are taken in every location. Therefore, in a developing world like Ghana for example, its implementation may not be possible, at least for now. In this respect, an alternative weighting system, which is comparatively cheaper and makes it possible to compute weights for both base and current year price data, is suggested in this paper.

In the search for an alternative, it is important to recognise that the essence of weighting in indexing is always linked to variability (Hill, 2004; International Labour Office, & Turvey, 2004; Sellwood, 1989; Spiegel & Stephens, 2011; Srivastava & Srivastava, 2003) and that price datasets are usually multivariate in nature. However, current weighting system uses univariate approach only, denying the analysis of interaction among variables and the only time a multivariate method was applied to price data was by (Mettle et al., 2014), but that was not about weighting. Thus, weights can be obtained for price variables by subjecting the same price data to factor analysis and generate values that serve the same purpose as weights. Once a correlation matrix:

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

$$\boldsymbol{R} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & 1 \end{pmatrix}, \tag{7}$$

can be formulated for some price variables, $X_1, X_2, \dots, X_p$, a spectral decomposition of $\boldsymbol{R}$ allows for the variability accounted in each variable to be estimated. Thus $\boldsymbol{R}$ reveals factors common to all variables in the data that help in estimating a certain percentage of variability attributable to each variable. Essentially, factor analysis can produce weights. Detail methods and procedures are provided in the next section.

METHODOLOGY

The methodology involved in factor analysis can be utilised to build CPI weighting system that can serve as an alternative to the current system. According to literature on multivariate statistics, for example, in Johnson & Wichern (2007); Rencher (2002); and Timm (2002), if the variables: $X_1, X_2, \dots, X_p$ account for the maximum variance in a dataset, then some $m < p$ constructs will account for the inter-correlations among the original variables. So that, the factor model suggests that Xs are linearly dependent upon a few unobservable constructs (also called factors), $F_1, F_2, \dots, F_m$, common to all variables; and additional sources of variation, $\epsilon_1, \epsilon_2, \dots, \epsilon_p$, specific to the original variables in the model in the manner:

$$\begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{pmatrix} = \begin{pmatrix} l_{11} & l_{21} & \cdots & l_{p1} \\ l_{12} & l_{22} & \cdots & l_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ l_{1p} & l_{2p} & \dots & l_{pm} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_p \end{pmatrix}, \tag{8}$$

called the classical factor analysis model, where μ_i are the population means of the original variables and l_{ij} are simply, correlations between the original variables and the extracted factors, also called factor loadings. Frequently, (8) can be re-rewritten as:

$$\boldsymbol{X} = \boldsymbol{\mu} + \mathbf{L}\boldsymbol{F} + \boldsymbol{\epsilon}, \tag{9}$$

under the assumptions: (1) $\boldsymbol{F} \sim (\mathbf{0}, \mathbf{I}_m)$ and are uncorrelated; (2) $\boldsymbol{\epsilon} \sim (\mathbf{0}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi} = diag(\psi_1, \psi_2, \dots, \psi_p)$, diagonal matrix of specific variances and (3) for any pairs j and k, ϵ_j and F_k are independent. Subsequently, $Cov(\boldsymbol{\epsilon F}) = E(\boldsymbol{\epsilon F}^T) = \mathbf{0}$. From here, we can formulate the covariance structure of $\boldsymbol{X}$ as:

$$\begin{aligned} Var(\boldsymbol{X}) &= E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'] \\ &= E[(\mathbf{L}\boldsymbol{F} + \boldsymbol{\epsilon})(\mathbf{L}\boldsymbol{F} + \boldsymbol{\epsilon})'] \\ &= E[\mathbf{L}\boldsymbol{F}\boldsymbol{F}'\mathbf{L}'] + E[\mathbf{L}\boldsymbol{F}\boldsymbol{\epsilon}'] + E[\boldsymbol{\epsilon}\boldsymbol{F}'\mathbf{L}'] + E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}'] \\ &= \mathbf{L}E[\boldsymbol{F}\boldsymbol{F}']\mathbf{L}' + \mathbf{L}E[\boldsymbol{F}\boldsymbol{\epsilon}'] + E[\boldsymbol{\epsilon}\boldsymbol{F}']\mathbf{L}' + E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}'] \end{aligned}$$

$$\therefore \boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}' + \boldsymbol{\Psi}. \tag{10}$$

So that, the covariance between $\boldsymbol{X}$ and $\boldsymbol{F}$ takes the form:

$$\begin{aligned} Cov(\boldsymbol{X}, \boldsymbol{F}) &= E[(\boldsymbol{X} - \boldsymbol{\mu})\boldsymbol{F}'] \\ &= E[(\mathbf{L}\boldsymbol{F} + \boldsymbol{\epsilon})\boldsymbol{F}'] \\ &= E[\mathbf{L}\boldsymbol{F}\boldsymbol{F}'] + E[\boldsymbol{\epsilon}\boldsymbol{F}'] = \mathbf{L}, \end{aligned}$$

implying

$$Cov(X_i, f_j) = l_{ij}\,. \tag{11}$$

Also from (10), and for any number of variables, p, we have

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{pmatrix} =$$

$$\begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1} & l_{p2} & \dots & l_{pm} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & \cdots & l_{p1} \\ l_{12} & l_{22} & \cdots & l_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ l_{1m} & l_{2m} & \dots & l_{pm} \end{pmatrix} + \begin{pmatrix} \psi_1^2 & 0 & \dots & 0 \\ 0 & \psi_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \psi_p^2 \end{pmatrix}.$$

So that, performing matrix multiplication operation,

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{m} l_{1j}^2 & \sum_{k=1}^{m} l_{1k}l_{2k} & \dots & \sum_{k=1}^{m} l_{1k}l_{pk} \\ \sum_{k=1}^{m} l_{2k}l_{1k} & \sum_{j=1}^{m} l_{2j}^2 & \dots & \sum_{k=1}^{m} l_{2k}l_{pk} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} l_{pk}l_{1k} & \sum_{k=1}^{m} l_{pk}l_{2k} & \dots & \sum_{j=1}^{m} l_{pj}^2 \end{pmatrix}$$

$$+ \begin{pmatrix} \psi_1^2 & 0 & \dots & 0 \\ 0 & \psi_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \psi_p^2 \end{pmatrix}.$$

$$\Rightarrow \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{m} l_{1j}^2 + \psi_1^2 & \sum_{k=1}^{m} l_{1k}l_{2k} & \dots & \sum_{k=1}^{m} l_{1k}l_{pk} \\ \sum_{k=1}^{m} l_{2k}l_{1k} & \sum_{j=1}^{m} l_{2j}^2 + \psi_2^2 & \dots & \sum_{k=1}^{m} l_{2k}l_{pk} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} l_{pk}l_{1k} & \sum_{k=1}^{m} l_{pk}l_{2k} & \dots & \sum_{j=1}^{m} l_{pj}^2 + \psi_p^2 \end{pmatrix}.$$

Equating the terms yield:

$$\sigma_i^2 = \sum_{j=1}^{m} l_{ij}^2 + \psi_i^2 = \sum_{j=1}^{m} l_{ij}^2 + Var(\varepsilon_i), \qquad for\ i = 1,2,\dots,p \tag{12}$$

and,

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

$$\sigma_{ij} = Cov(X_i, X_j) = \sum_{k=1}^{m} l_{ik} l_{jk}, \qquad for\ i, j = 1,2, \dots, p. \tag{13}$$

Note that from (10), the parameters of interest are $\mathbf{L}$ and $\mathbf{\Psi}$. To estimate these, principal component is one of the methods frequently used. The principal component method of factor estimation allows all variables to have an initial communality of one (1) and seeks to estimate the final communality by spectral decomposition of the correlation matrix, (7). A spectral decomposition expresses the correlation matrix in terms of its eigenvalues and eigenvectors. Assume the eigenvalues of (7) are such that: $\lambda_1 > \lambda_2 > \cdots > \lambda_p$; and the corresponding set of eigenvectors: $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_p$; then we can rewrite $\mathbf{R}$ as:

$$\mathbf{R} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1' + \lambda_2 \mathbf{e}_2 \mathbf{e}_2' + \cdots + \lambda_p \mathbf{e}_\mathrm{p} \mathbf{e}_\mathrm{p}',$$

so that

$$\mathbf{R} = \left(\boldsymbol{e}_1\sqrt{\lambda_1}, \boldsymbol{e}_2\sqrt{\lambda_2}, \dots, \boldsymbol{e}_p\sqrt{\lambda_p}\right) \begin{pmatrix} \boldsymbol{e}_1'\sqrt{\lambda_1} \\ \boldsymbol{e}_2'\sqrt{\lambda_2} \\ \vdots \\ \boldsymbol{e}_p'\sqrt{\lambda_p} \end{pmatrix}. \tag{14}$$

Clearly, (14) is in the form of (10) in the case where all p variables are extracted. When that happens, (10) reduces to $\mathbf{LL}'$ implying that, $Cov(X_i, f_j) = \boldsymbol{e}_i\sqrt{\lambda_i}$, called the unrotated factor matrix. Even though the unrotated initial factor loadings have the ability to reproduce the standardized covariance matrix, it does pose interpretability problems (Rencher, 2002). So, to obtain a simple factor structure that provides a most meaningful and interpretable set of factors and help to reduce the chances of having bipolar factors, varimax rotation is adopted (Conway & HufFStutt, 2003; Reimann et al., 2002; Rencher, 2002). Let $\mathbf{T}$ be an $m \times m$ orthogonal transformation matrix such that $\mathbf{TT}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$. Then the product, $\mathbf{LT}$ is a rotated factor loading matrix of $\mathbf{L}$ under $\mathbf{T}$. Therefore, we can define a new rotated matrix, $\mathbf{L}^* = \mathbf{LT}$, such that from (10), $\mathbf{\Sigma} = \mathbf{LTT}'\mathbf{L}' + \mathbf{\Psi} = (\mathbf{LT})(\mathbf{LT})' + \mathbf{\Psi}$.
So that,

$$\mathbf{\Sigma} = \mathbf{L}^*\mathbf{L}^{*\prime} + \mathbf{\Psi}. \tag{15}$$

Thus, (15) is the final rotated factor model. At this point, it serves intuitive purpose to determine the amount of variability in each variable accounted for by the m extracted factors, which is the communality. This variability is part of (12). Equation (12) breaks down the total variance in a variable to comprise the communality and the specific variance: $\sigma_i^2 = \sum_{j=1}^{m} l_{ij}^2 + Var(\varepsilon_i)$, hence the i^{th} communality:

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + \cdots + l_{im}^2. \tag{16}$$

Equation (16) is the communality obtained for a classical factor analysis; it accounts for the variability in each variable and hence fit to be used as weights on variables. When used as weights,

$$w_i^* = \frac{h_i^2}{\sum_{i=1}^{p} h_i^2} \tag{17}$$

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

its additional advantage is that it retains the variability in each price variable in the final index so that the final index fairly represents variability in each price item. This is the proposed multivariate weighting system, w_i^*, which can be found for both base and current year datasets in real time, once price datasets are obtained.

However, before arriving at (16), two important issues need to be cleared. The first is the criteria for retaining factors. A more popular method for factor retention is the eigenvalue-greater-than-one rule. However, a relatively recent method: parallel analysis, is a superior and robust alternative (Basto & Pereira, 2012; Beaujean, 2013; Ledesma & Valero-Mora, 2007). For this reason, both methods were employed in this study and results compared to ascertain the effect of different factor retention method when using the multivariate weighting system proposed in this study. The second has to do with why the choice of principal component method of factor estimation over other methods of factor estimation. Clearly, the essence of weighting in indexing is to incorporate variability of variables in the final index. So, to use communality as weights, a good amount of common variance in the correlation matrix ought to have been account for. That is exactly what the principal component method helps to achieve. For a large set of price variables, the eigenvalues of the correlation matrix partitions the total common variance. So that, the variable(s) with eigenvalues greater than one would explain maximum variance in the correlation matrix. Therefore, for m extracted factors, percentage explained variance (PEV):

$$PEV = \left(\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_m}{p}\right)100, \tag{18}$$

would mostly be maximum. Again, it is only the principal component method of factor estimation that analyses the correlation matrix, other methods perform analysis with a reduced correlation matrix. In addition, for relatively large number of variables (which may be the case mostly in price indexing), principal component analysis is an approximation of exploratory factor analysis (Gorsuch, 1986; Lorenzo-Seva, 2013; Snook & Gorsuch, 1989; Thompson, 1992). By this, it should be the case that (18) would always account for the maximum common variance in the dataset. Thus, the classical multivariate weighting system for consumer price index is proposed as in (17).

To test the proposed method and compare results to the current method, both simulation and empirical studies were carried out. In the simulation studies pseudo price and weight datasets are required. In this regard, while price values were assumed to follow normality, weights datasets were made to follow the uniform distribution with end points, 0 and 1, since current weights are within these values. The different conditions considered in the simulation studies are: (1) to investigate the effects of the two weighting systems for different sets of variables; (2) to investigate the effects of the two weighting systems for different sample sizes and (3) to investigate the effects of the two weighting systems for both high and low KMO values. To investigate the different effects of the two weighting systems, the multivariate location and scatter for the following price datasets were used to simulate pseudo price datasets: (1) a 19-variable datasets from the Ghana Statistical Service (GSS) and (2) a 55-variables datasets obtained from the internet (specifically, www.numbeo.com), with simulated weights values as described earlier. For each set, different sample sizes were simulated aimed

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

at investigating the power of the Bartlett's sphericity test and KMO. Literature, for example in Budaev (2010); Dziuban et al. (1979); Knapp & Swoyer (1967); and Wilson & Martin (1983) proposed sample sizes of $N = 50, 100, 250, 500$ and 1000 in studying the behaviour of Bartlett's sphericity test and KMO, which are the two key tests in performing factor analysis that satisfy the assumptions stated earlier.

For the empirical study, two immediate points in time where HES was conducted in a country are needed. Accordingly, 2002 and 2012 data on Household Expenditure Survey (HES) in Ghana, produced by the Ghana Statistical Service (GSS), were used as base and current years respectively. Average prices were then computed for each region for a year, amounting to 120 cases each, since there were 10 regions in Ghana at the time of carrying out this study. Therefore, the maximum number of price variables required for factor analysis in this case is 24. Twenty-four price items were therefore randomly selected from the basket of 267 items; and out of the 24, nineteen (19) items were found to be common for both 2002 and 2012. Table 1 presents the selected items together with their rescaled expenditure-based weights as obtained from GSS.

Table 1: Empirical Price Data Variables with Their Corresponding Weights

			Expenditure-Based Weights	
S/N	Random Number	Item Name and Notation	2002	2012
1	26	X_1=Snails	5.42	9.10
2	90	X_2=Dark-Beer	2.97	2.69
3	234	X_3=Tuition Fees	2.13	7.32
4	31	X_4=Tuna	13.65	9.10
5	129	X_5=Women Sandals	5.02	2.89
6	200	X_6=Wheel Alignment	4.13	0.30
7	179	X_7= ATC Drug	5.77	2.88
8	121	X_8=Handkerchief	5.02	0.25
9	43	X_9=Coconut (fresh)	2.83	3.41
10	34	X_{10}=Evaporated Milk	2.25	3.44
11	205	X_{11}=Taxi Fare	4.13	7.34
12	13	X_{12}=Sugar Bread	5.32	6.77
13	155	X_{13}=Refrigerator	10.43	0.74
14	56	X_{14}=Tomatoes (fresh)	5.53	9.29
15	10	X_{15}=Cassava Dough	5.53	6.77
16	252	X_{16}=Toothpaste	4.00	4.97
17	2	X_{17}= Imported Rice	5.32	6.77
18	102	X_{18}=T-shirt	5.02	6.68
19	52	X_{19}=Garden-eggs	5.53	9.29

Source: Ghana Statistical Service (GSS)

The procedures for data analysis are the same for both simulation and empirical studies: (i) For each sample price data, made up of both base and current year datasets,

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

multivariate weights are determined by the classical factor analysis method; (ii) Price relativity data is found by dividing current year prices with corresponding base year prices; (iii) Expenditure-based weights are those collected from GSS or their pseudo values, in the case of simulation studies; (iv) The two set of weights (multivariate and expenditure-based weights) are then substituted into the standard index formulas considered for this study, which are functions of the price relativity data; (v) Procedures (i) to (iv) are repeated for one thousand 1000 bootstrap samples; (vi) Then the proportion of all 1000 indices that fall within expenditure-based Laspeyres' index and expenditure-based Paasche's index (the range believed to contain the ideal index) are determined. Standard errors of bootstrap estimates for the two weighting systems were also compared. These procedures were carried out for all different conditions and different datasets for both empirical and simulation studies. Detail discussions on procedures (v) and (vi) are presented next.

Ordinarily, a best and efficient estimator is one with lower standard error. However, because the distribution of indices by the two set of weighting systems are not known, the theoretical formulas for computing standard error in this case cannot not be determined. Therefore, bootstrapping would be the appropriate method to estimate the standard errors for comparison. In this case, the bootstrap algorithm for estimating the standard error of an estimator would draw 1000 samples with replacement from the original sample data matrix, $X_{ij}, i = 1,2, \dots, n; j = 1,2, \dots, p$ with each data case having the same probability, $\frac{1}{n}$, of being drawn; so that, each sampled data matrix is denoted as, $X^*_{(ij)k}$, for all $k = 1,2, \dots, 1000$ re-samples. In general, if θ is an unbiased estimator for a given data population, then for pairs of unbiased estimators for θ, $(\hat{\theta}_i, \hat{\theta}_j)$, where $i, j = 1,2, \dots, k$, the efficiency of $\hat{\theta}_i$ relative to $\hat{\theta}_j$ is the ratio:

$$e\left(\hat{\theta}_i, \hat{\theta}_j\right) = \frac{var(\hat{\theta}_j)}{var(\hat{\theta}_i)}. \qquad (19)$$

That is, if $e\left(\hat{\theta}_i, \hat{\theta}_j\right) > 1$ or $var\left(\hat{\theta}_j\right) > var(\hat{\theta}_i)$, then $\hat{\theta}_i$ is relatively more efficient than $\hat{\theta}_j$. Thus (19) is important in determining the efficiency of an index estimator with multivariate weights and that for index estimators with expenditure-based weights. Now, for each bootstrap sample, $X^*_{(ij)k}$, eight indexes would be found: the first-four would be having expenditure-based weights substituted into (1), (2), (3) and (4); while the last-four would have multivariate weights substituted in same. So, for the $t = 1,2, \dots, 8$ index estimators, overall mean for all cases and grand mean would be determined respectively for all N bootstrap samples as:

$$\bar{I}^*_{(t)k} = \frac{1}{n}\sum_{i=1}^{120} I_{(t)i}\,; \quad \bar{I}^{**}_t = \frac{1}{1000}\sum_{k=1}^{1000} \bar{I}^*_{(t)k}. \qquad (20)$$

From (20), the bootstrap estimate of variance and standard errors are respectively found as:

$$\hat{\sigma}^2_{(bs)t} = \frac{1}{1000-1}\sum_{k=1}^{1000}\left(\bar{I}^{*}_{(t)k} - \bar{I}^{**}_{t}\right)^2; \quad \sqrt{\hat{\sigma}^2_{(bs)t}}. \tag{21}$$

In making comparisons, estimates with least standard error are relatively most efficient. Finally, the percentage of the 1000 bootstrap estimates that fall between the expenditure-weighted Laspeyres' and Paasche's index values would be determined too. This range is believed to hold superlative or ideal indices. To do this, assume n_i is the number of times an estimator produces grand index estimates that lie between the expenditure-based Laspeyres' index, I_1, and the expenditure-based Paasche's index, I_4, in all the 1000 bootstrap samples. Then, the percentage of times estimates fall in the ideal index range ($\%LP_{range}$) is:

$$\%LP_{range} = \frac{n_i}{1000} \times 100, \tag{22}$$

where n_i $(i = 1, 2, \ldots, 8)$ are the number of times the i^{th} index number falls within the ideal range.

RESULTS

Results are reported for both simulation and empirical studies. In this regard, outputs relevant for making inference are: (1) the KMO value and Bartlett's significance value of each price dataset; (2) bootstrap mean estimates for both expenditure-based consumer price indices (ebCPI) and multivariate-based consumer price indices (mwCPI); (3) standard errors for both ebCPI and mwCPIs; and (4) the proportion of all 1000 bootstrap estimates that fall within the expenditure-based Laspeyres' index and expenditure-based Paasche's index. In addition, the empirical study would sought to investigate if the two factor retention methods – eigenvalue-greater-than-one method or the parallel analysis method would produce different results. Tables 2 and 3 present results of the simulation study while Table 4 presents results of the empirical study. Subsequently, the following deductions can be made, based on the simulation results:

(i) In all cases, standard error induced by multivariate weights is lowest, showing that multivariate weights are most efficient in estimating indices.

(ii) Majority of cases (7 out of 10 representing 70%) multivariate weights produced indices falling within the ideal index range.

(iii) In two out of the ten cases (20%), the two performed equally; while the rest 10% of the cases, the expenditure-based weighting system edged-past the proposed weighting system.

(iv) There is no evidence to suggest that the findings made in (i) to (iii) above are different for different number of variables involved in the analysis.

(v) There is also no evidence to suggest that the findings in (i) to (iii) above are different for low or high KMO, or for varying sample sizes.

In a nutshell, results from the simulation studies have shown that irrespective of the number of variables, sample size, and KMO value, index estimates produced by the expenditure-based weights in standard index formulas results in higher standard errors than those produced with the multivariate-based weights in same formulas. Also, in about 90%(70%+20%) of the cases considered in this study, multivariate weighted CPIs mimicked or out-performed the expenditure-based CPIs in terms of proportion of indices falling within the ideal index range. The following deductions can be made from results of the empirical studies, as depicted in Table 4:

(i) In both cases, standard errors induced by the expenditure-based weights are similar to those induced by multivariate weights; standard errors round up to same value.

(ii) Results for multivariate-weighted CPIs via eigenvalue-greater-than-one method are the same for those via parallel analysis method.

(iii) On the proportion of estimates falling within the ideal index range, multivariate-weighted consumer price indices out-performed the expenditure-based consumer price for Laspeyre's and Paasche's indices.

(iv) For the superlative indices, expenditure-based consumer price index edged passed multivariate-weighted consumer price index slightly: the differences in the neighbourhood of 0.02.

Clearly, the findings of both simulation and empirical studies appear to have set out the proposed multivariate weighting system as a good alternative to the current expenditure-based weighting system.

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

	KMO (Bartlett's sig.)			ebCPI		mwCPI		Proportion Within Ideal Index Range	
Sample Size	Base Year	Current Year	Index	Bootstrap Estimate	Standard Error	Bootstrap Estimate	Standard Error	ebCPI	mwCPI
50	0.6(0.00)	0.4(0.00)	Laspeyre's	6.05	0.77	5.56	0.16	0.35	0.96
			Paasche's	5.32	0.82	5.58	0.16	0.32	0.94
			Fishers'	5.67	0.57	5.56	0.15	0.47	0.95
			Drob.Bowley	5.69	0.57	5.57	0.15	0.48	0.95
100	0.6(0.00)	0.4(0.00)	Laspeyre's	6.24	1.84	7.65	1.17	0.43	0.57
			Paasche's	8.13	1.91	7.70	1.21	0.43	0.56
			Fishers'	6.77	1.50	7.67	1.18	0.50	0.57
			Drob.Bowley	7.18	1.54	7.68	1.19	0.50	0.57
250	0.6(0.00)	0.5(0.00)	Laspeyre's	7.42	0.98	6.12	0.19	0.87	1.00
			Paasche's	4.56	0.99	6.11	0.19	0.85	1.00
			Fishers'	5.78	0.70	6.11	0.19	0.97	1.00
			Drob.Bowley	5.99	0.70	6.12	0.19	0.96	1.00
500	0.6(0.00)	0.5(0.00)	Laspeyre's	6.88	1.08	6.73	0.41	0.06	0.18
			Paasche's	6.71	1.14	6.81	0.42	0.06	0.15
			Fishers'	6.77	0.85	6.76	0.41	0.08	0.17
			Drob.Bowley	6.80	0.84	6.77	0.41	0.08	0.17
1000	0.6(0.00)	0.5(0.00)	Laspeyre's	7.78	1.07	6.75	0.27	0.13	0.03
			Paasche's	7.29	1.09	6.73	0.26	0.15	0.02
			Fishers'	7.48	0.79	6.74	0.26	0.12	0.03
			Drob.Bowley	7.54	0.78	6.74	0.26	0.14	0.03

Table 2: Simulation Results of Multivariate-Weighted CPI and Expenditure-Based CPI for $p = 19$

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

Source: Statistical Analysis of Simulated Datasets; p is the number of variables

Table 3: Simulation Results of Multivariate-Weighted CPI and Expenditure-Based CPI for $p = 55$

	KMO (Bartlett's sig.)			ebCPI		mwCPI		Proportion Within Ideal Index Range	
Sample Size	Base Year	Current Year	Index	Bootstrap Estimate	Standard Error	Bootstrap Estimate	Standard Error	ebCPI	mwCPI
50	0.5(0.00)	0.5(0.00)	Laspeyre's	14.53	3.13	12.53	1.74	0.02	0.03
			Paasche's	14.27	2.98	12.81	1.74	0.03	0.03
			Fishers'	14.34	2.48	12.67	1.73	0.03	0.03
			Drob.Bowley	14.40	2.50	12.67	1.73	0.03	0.03
100	0.8(0.00)	0.9(0.00)	Laspeyre's	12.00	6.74	18.42	3.40	0.84	0.99
			Paasche's	29.63	6.95	19.62	3.49	0.82	0.99
			Fishers'	16.54	5.61	19.00	3.41	0.91	0.99
			Drob.Bowley	20.79	5.58	19.02	3.42	0.92	0.99
250	0.9(0.00)	0.9(0.00)	Laspeyre's	17.63	2.97	13.63	1.43	0.75	0.98
			Paasche's	10.63	2.98	14.44	1.39	0.74	0.99
			Fishers'	13.06	2.39	14.02	1.38	0.86	0.99
			Drob.Bowley	14.13	2.38	14.03	1.38	0.86	0.99
500	0.9(0.00)	0.9(0.00)	Laspeyre's	11.60	3.52	15.61	1.70	0.54	0.88
			Paasche's	17.73	3.44	15.84	1.72	0.57	0.84
			Fishers'	13.71	2.79	15.72	1.68	0.69	0.88
			Drob.Bowley	14.67	2.80	15.72	1.69	0.62	0.87

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

Imputed Data	KMO (Bartlett's sig.) Base Year	Current Year	Index	ebCPI Bootstrap Estimate	ebCPI Standard Error	mwCPI Bootstrap Estimate	mwCPI Standard Error	Proportion Within Ideal Index Range ebCPI	Proportion Within Ideal Index Range mwCPI
Eigenvalue Method	0.5(0.00)	0.6(0.00)	Laspeyre's	4.65	0.07	5.59	0.11	0.37	0.99
			Paasche's	5.81	0.09	5.71	0.11	0.50	0.93
			Fishers'	5.18	0.07	5.65	0.11	1.00	0.98
			Drob.Bowley	5.23	0.07	5.65	0.11	1.00	0.98
Parallel Analysis Method	0.5(0.00)	0.6(0.00)	Laspeyre's	4.65	0.07	5.59	0.12	0.37	0.98
			Paasche's	5.81	0.09	5.71	0.11	0.50	0.92
			Fishers'	5.18	0.07	5.65	0.10	1.00	0.98
			Drob.Bowley	5.23	0.07	5.65	0.10	1.00	0.98
1000	0.9(0.00)	0.9(0.00)	Laspeyre's	33.77	10.94	31.37	8.94	0.003	0.004
			Paasche's	33.88	11.00	31.95	9.09	0.004	0.005
			Fishers'	31.96	9.16	31.66	9.01	0.003	0.006
			Drob.Bowley	33.82	9.63	31.66	9.01	0.004	0.007

Source: Statistical Analysis of Simulated Datasets; p is the number of variables

Table 4: Empirical Results of Multivariate-Weighted CPI and Expenditure-Based CPI

Source: Statistical Analysis of Simulated Datasets

International Journal of Mathematics and Statistics Studies
Vol.8, No.3, pp.1-19, October 2020
Published by ***ECRTD-UK***
Print ISSN: 2053-2229 (Print)
Online ISSN: 2053-2210 (Online)

DISCUSSIONS

Ideally, discussions require comparison of results and findings of a research to other results in literature. However, in this research, there are no results and findings regarding the subject. Clearly, there are no counterpart results and findings except those of this study, since communality has not yet been proposed as an alternative weighting scheme for consumer price index formulation and there are no alternative index weighting systems proposed in literature either. Therefore, the focus of the discussion under this section would be to discuss the ramifications of results and findings on the body of knowledge in index numbers and consumer price indexing.

Currently, there are at least four standard weighted index formulas used world-wide for consumer price indexing. These are: (1) Laspeyres' index, which uses weights from base year; (2) Paasche's index, which uses weights from a current year; (3) Fisher's Ideal index, which is the geometric mean of Laspeyres' and Paasche's indices and (4) Drobish-Bowley's index, an arithmetic mean of Laspeyres' and Paasche's indices. Weighted price indices are preferred because they allow for the relative importance or relevance of individual price items to be incorporated in the resultant index. This is why the importance of weighting in price indexing cannot be overemphasized in the whole essence of index numbers.

This notwithstanding, the problem of index numbers and by extension consumer price index, has been weight and formula biases that are largely because expenditure-based weights are difficult or expensive to generate for both base and current years and used immediately for indexing. In view of this, it has become virtually impossible to formulate superlative indices, which are deemed to be ideal. Implicitly, the formula bias is a consequence of the weight bias. The implication is that if weights can be obtained for both base and current years anytime indices are to be composed, both problems will cease to exist.

So logically, the immediate solution to both weighting and formula biases is to have an alternative weighting system that activates the use of both Laspeyres' and Paasche's indices, which then makes it possible to compute superlative indices like the Fisher's Ideal and Drobish-Bowley's indices. This has been the main objective of this study.

This study presented a multivariate weighting system as an alternative weighting system in index formulations. Here, it is important to state that the multivariate weighting system, which utilises communality to formulate weights, does not erode the essence of weighting as already known in literature regarding indexing, for example in Hill (2004); International Labour Office & Turvey (2004); Sellwood (1989); Spiegel & Stephens (2011) and in Srivastava & Srivastava (2003). In fact, since communality measures the amount of variability in variables explained by some common factors, it serves the same purpose indexing with communality as an alternative weighting system. Essentially, communality as weights retains the explained variability of each variable in the resultant index.

The proposed methodology was tested using simulation studies and later using real data collected from the Ghana Statistical Service (GSS). In both cases, results for the multivariate approach to weighting were compared to those of the current univariate method. In the simulation studies, the new methodology was subjected to different conditions precedent in the conduct of factor analysis. Two different datasets recording various KMOs, ranging from low to high, were generated under the assumption of multivariate normality. For each set of data, five different sample sizes popularly known in literature for checking the behaviour of KMO and Bartlett's test of sphericity, were adopted. Ultimately, the effect of multivariate weights in standard index numbers composed via classical factor analyses and the two factor retention methods, eigenvalue-greater-than-one rule and parallel analysis, were assessed.

Results from the simulation studies have shown that irrespective of the number of variables, sample size and KMO value, index estimates produced by the expenditure-based weights in standard index formulas resulted in higher standard errors than those produced with the multivariate-based weights in same formulas. Also, in 70% of the cases considered, multivariate weighted CPIs clearly out-performed the expenditure-based CPIs in terms of proportion of indices falling within the ideal index range. Here again, the standard errors of indices produced with multivariate weights were lesser than those produced with pseudo expenditure-based weights. Clearly, these findings appeared to have set out the proposed multivariate weighting system as a good alternative to the current expenditure-based weighting system.

The empirical study was required in order to test the new theory on real data and to ascertain the findings in the simulation studies. To do this, there was the need to select two points in the past (2002 and 2012) where household expenditure data were generated in Ghana and use them as base and current years respectively. This is to help obtain the ideal index range in this case. Nineteen (19) items were included.

Subsequently, results of the empirical study, after bootstrapping a thousand (1000) times for each set, confirmed that the two weights in standard index formulas are equally efficient and standard errors for bootstrap estimates were approximately the same for all cases considered. Again in most cases, the multivariate-weighted CPIs outperformed the expenditure-based CPIs in terms of indices that fell between the ideal index range. In the other few cases where the expenditures-based CPIs edged past the multivariate-weighted CPIs, the difference was in the neighbourhood of 0.02. This suggests that the multivariate-weighted CPIs do mimic expenditure-based weights closely in producing indices that would be deemed as ideal. In this regard, the multivariate-weighted consumer price index values have shown to be consistent with the theory on the ideal index, as found in many literatures; for example, in Afriat (2005 &1977); Hicks (1956); Konus (1924); Samuelson & Swamy (1974); Samuelson (1974 & 1984) and in Swamy (1984), since its values are closer to the Fisher's and Drobish-Bowley's indices.

The implication is that it should be possible to generate multivariate weights conveniently from price data for used in the composition of CPI without having to conduct expensive household expenditure survey any time indices are to be composed. This will not only solve the main problem of index numbers, but will, also, allow for other index values to be computed in real time, once price data are loaded onto a platform. Essentially, the weighting system proposed in this study can help solve the long standing weight and formula problems associated with CPI.

Implication to Research and Practice

1. The proposed weighting system is recommended for use in the formulation of the consumer price index.
2. The proposed weighting system can also be applied in the financial sector and in performance indexing.

CONCLUSIONS

The study sought to propose a multivariate weighting system for CPI by applying classical factor analysis methodology to mitigate the widely reported weight and formula biases, with additional advantage of explaining variability in price variables. Evidently, this main objective was achieved. Consequently, the following conclusions and summaries can be made:

1. Multivariate weighted CPIs recorded lesser standard errors than expenditure-based CPIs in the simulation studies, hence appears to be most efficient.
2. In the practical case, standard errors of CPIs composed with the two weighting systems are approximately the same, suggesting that multivariate weighted system can be a good alternative to the expenditure-based weighting system.
3. In 90% of all cases considered in this study, multivariate weighted CPIs mimicked or out-performed the expenditure-based CPIs in terms of proportion of indices falling within the ideal index range.
4. Again in the practical case, multivariate weighted CPIs had higher proportions of indices within the ideal index range than the expenditure-based CPIs.
5. Multivariate-weighted consumer price indices for different factor retention methods are equally efficient.
6. There was no evidence that the conclusion above differ for different number of variables, sample sizes, and KMO values.
7. It can be asserted that the proposed method will work for all datasets since it performed comparably better for both the empirical and simulated datasets; hence it will be a judicious alternative for the current CPI weighting system.

Future Research

The methodology of adopted in this paper made use of classical factor analysis under the assumption that the price data is normally distributed. However, this assumption of normality may not always hold for all multivariate price datasets. In such situations, the robust counterpart of factor analysis would be appropriate for the generating weights for the price variables. A future research will focus on this aspect.

References

Afriat, S. N. (2005). *The price index and its extension: a chapter in economic measurement*. London and New York: Routledge.

Afriat, S. N. (1977). *The price index*. Cambridge, UK: Cambridge University Press.

Basto, M., & Pereira, J. M. (2012). An SPSS R-Menu for ordinary factor analysis. *Journal of Statistical Software*, *46*(4). Retrieved from www.jstatsoft.org

Beaujean, A. A. (2013). Factor analysis using R. *Practical Assessment, Research and*

Evaluation, *18*(4). Retrieved from www.pareonline.net/getvn.asp?v=18andn=4

Boskin, M.J.,Dulberger, E.R., Gordon, R.J., Griliches, Z., Jorgenson, D. W. (1996). *Towards a more accurate measure of the cost of living*. Final Report, Senate Finance Committee.

Bryan, M. F., & Cecchetti, S. G. (1993). The consumer price index as a measure of inflation. *National Bureau of Economic Research(NBER),Working Paper, No.4505*.

Budaev, S. (2010). *Multivariate methods and small sample size: combining with small effect size*. Available at SSRN 3075441.

Cecchetti, S. G. (1996). Measuring short-run inflation for central bankers. *National Bureau of Economic Research(NBER),Working Paper*, *No.5786*.

Clark, T. E. (1999). A comparison of the CPI and the PCE Price Index. *Economic Review-Federal Reserve Bank of Kansas City*, *84*(3), 15.

Conway, J. M., & HufFStutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organisational research. *Organizational Research Methods*, *6*(2), 147–168.

Dziuban, C.D., Shirkey, E.S., & Peeples, T. O. (1979). An investigation of some distributional characteristics of the measure of sampling adequacy. *Educational Psychological Measure*, *39*, 543–549.

Gorsuch, R. . (1986). *Factor analysis* (2nd ed). Hillside, New York, NY: Erlbaum.

Hicks, J. . (1956). *A revision of demand theory*. Oxford: The Clarendon Press.

Hill, P. (2004). Consumer price index manual. *Statistical Journal of the United Nations Economic Commission for Europe*, *21*(2), 87–93.

International Labour Office, & Turvey, R. (2004). *Consumer price index manual: theory and practice*. International Monetary Fund.

Johnson, R.A., Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed.). Englewood Cliffs, New Jersey: Prentice – Hall International, Inc.

Khalid, M., Asghar, Z. (2010). *Working to improve price indices development in Pakistan* (PIDE Working Papers No. 62). Islamabad.

Knapp, T.R., & Swoyer, V. H. (1967). Some empirical results concerning the power of Bartlett's test of the significance of a correlation matrix. *American Educational Research Journal*, *4*, 13–17.

Konus, A. . (1924). The problem of the true index of the cost of living. *Econometrika*, *7*(Jan. 1939), 10–29.

Ledesma, R. D., and Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment Research and Evaluation*, *12*(2).

Lorenzo-Seva, U. (2013). *How to report the percentage of explained common variance in expploratory factor analysis*. Technical Report. Department of Psychology, Universitat Rovira I Virgili, Tarragona.

Malik, S.J., Nazli, H., Mehmood, A., and Shahzad, A. (2014). *Issues in the measurement and construction of the consumer price index in pakistan* (Working paper No. 020). International Food Policy Research Institute, Pakistan Strategy Support Program.

McCully, C.P., Moyer, B.C., & Steward, K. J. (2007). A reconciliation between the consumer price index and the personal consumption expenditures price index. *US Bureau of Economic Analysis and US Bureau of Labor Statistics, September, 12*.

Mettle, F.O., Yeboah, E., Asiedu, L. (2014). Profile analysis of spatial differential of inflation in Ghana. *International Journal of Statistics and Analysis*, *4*(2), 245–259.

Milana, C. (2009). *Solving the index-number problem in a historical perspective* (EU Klems Working Paper Series Product, nr.43).

Reed, S. B., and Rippy, D. A. (2012). Consumer price index data quality: how accurate is the U.S. CPI? In *Federal Publications, Cornell University's Digital Commons @ ILR School*.

Reimann, C., Filzmoser, P., and Garrett, R. G. (2002). Factor analysis applied to regional geochemical data: problems and possibilities. *Applied Geochemistry*, *17*(3), 185–206.

Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed.). New York, NY: John Wiley and Sons, Inc.

Samuelson, P.A., Swamy, S. (1974). Invariant economic index numbers and canonical duality: survey and synthesis. *American Economic Review*, *64*(4), 566–593.

Samuelson, P. A. (1974). Analytical notes on international real-income measures. *Economic Journal*, *84*(335), 595–608.

Samuelson, P. A. (1984). Second thoughts on analytical income comparisons. *Economic Journal*, *94*, 267–278.

Sellwood, D. J. (1989). *Consumer price indices: an ILO manual*. International Labour Office.

Snook, S.C., & Gorsuch, R. . (1989). Component analysis versus common factor analysis: A Monte Carlo Study. *Psychological Bulletin*, *106*, 148–154.

Spiegel, M.R., and Stephens, L. J. (2011). *Statistics*. New York, United States of America: McGraw-Hill book Company Limited.

Srivastava, S.C., and Srivastava, S. (2003). *Fundamentals of statistics* (1st ed.). New Delhi, India: Anmol Publications PVT. Ltd.

Swamy, S. (1984). Samuelson's analytical notes on international real-income measures: a comment. *Economic Journal*, *94*, 265–266.

Thompson, B. (1992). A partial test distribution for cosines among factors across samples. *Advances in Social Science Methodology, 2*, 81–97.

Timm, N. H. (2002). *Applied multivariate analysis*. New York, NY: Springger-Verlag, New York, Inc.

Wilson, G.A., & Martin, S. A. (1983). An empirical comparison of two methods for testing the significance of a correlation matrix. *Educational Psychological Measure*, *43*, 11–14.