STATISTICS IN TRANSITION-new series, Winter 2014
Vol. 15, No. 1, pp. 83–96

APPLICATION OF THE ORIGINAL PRICE INDEX FORMULA TO MEASURING THE CPI'S COMMODITY SUBSTITUTION BIAS

Jacek Białek[1]

ABSTRACT

This paper examines a possibility to apply the original price index formula to measuring the commodity substitution bias associated with the Consumer Price Index (CPI). Through simulation study the CPI bias values - calculated by using the original price index formula – is compared with those calculated on the basis of some known, superlative price indices.

Key words: CPI, COLI, superlative index, Laspeyres index, Fisher index.
JEL Classification: E17, E21, E30
AMS Classification: 62P20

1. Introduction

The Consumer Price Index (CPI) is used as a basic measure of inflation. The index approximates changes in the costs of household consumption that provide the constant utility (COLI, Cost of Living Index). In practice, the Laspeyres price index is used to measure the CPI (see White (1999), Clements and Izan (1987)). The Lapeyres formula does not take into account changes in the structure of consumption, which occur as a result of price changes in the given time interval. It means that the Laspeyres index can be biased due to the commodity substitution. Many economists consider the superlative indices (like the Fisher index or the Törnqvist index) to be the best approximation of COLI. Thus, the difference between the Laspeyres index and the superlative index should approximate the value of the commodity substitution bias. In this paper we propose the application of the original price index formula (see Białek (2012a), Białek (2013)) in measuring the commodity substitution bias associated with the Consumer Price Index (CPI). In our simulation study we compare the CPI bias values calculated by using the original price index formula with those calculated

[1] University of Lodz, Chair of Statistical Methods. E-mail: jbialek@uni.lodz.pl.

on the basis of some known, superlative price indices. It should be emphasized, that we do not consider other sources of the CPI biases, presented by White (1999).

2. Superlative price indices in the CPI bias measurement

Any discussion of consumer price index bias must first address the important issue of the target measure with respect to which the bias is measured. The final report of the Boskin Commission begins with a recommendation that "the Bureau of Labor Statistics (BLS) should establish a cost of living index (COLI) as its objective in measuring consumer prices" (see Boskin *et al.* (1996), page 2). Further discussions on the theory of the COLI can be found in the following papers: Diewert (1983), Jorgenson and Slesnick (1983), Pollak (1989).

Let $E(P,\bar{u}) = \min_Q \{P^T Q | U(Q) \geq \bar{u})$ be the expenditure function of a representative consumer which is dual to the utility function $U(Q)$. In other words it is the minimum expenditure necessary to achieve a reference level of utility $\bar{u}$ at vector of prices P. Then the Konüs cost of living price index is defined as

$$I_K = \frac{E(P^t,\bar{u})}{E(P^s,\bar{u})}, \tag{1}$$

where t denotes the current period, s denotes the base period, and in general, the vector of N considered prices at any moment τ is given by $P^\tau = [p_1^\tau, p_2^\tau, ..., p_N^\tau]^T$. I_K is a true cost of living index in which the commodity Q changes as the vector of prices facing the consumer changes. The CPI, in contrast, measures the change in the cost of purchasing a fixed basket of goods at a fixed sample of outlets over a time interval, *i.e.* $Q^s = [q_1^s, q_2^s, ..., q_N^s]^T = Q^t$.

The CPI is a Laspeyres-type index defined by

$$I_{La} = \frac{\sum_{i=1}^{N} q_i^s p_i^t}{\sum_{i=1}^{N} q_i^s p_i^s}, \tag{2}$$

so we assume here the constant consumption vector on the base period level. It can be shown (see Diewert (1993)) that under the assumption that the consumption vector Q^t solves the base period t expenditure minimization problem, that

$$I_K = \frac{E(P^t,U(Q^s))}{E(P^s,U(Q^s))} \leq I_{La}, \tag{3}$$

so $I_{La} - I_K$ is the extent of the commodity substitution bias, where I_K plays the role of the reference benchmark. In the so-called economic price index approach many authors use superlative price indices to approximate the I_K index (see White (1999)).

First we define a price index I to be *exact* for a linearly homogeneous aggregator function f (here a utility function), which has a dual unit cost function $c(P)$ and it holds

$$I = \frac{c(P^t)}{c(P^s)}. \tag{4}$$

In other words, an *exact* price index is one whose functional form is *exactly* equal to the ratio of cost functions for some underlying functional form representing preferences. The Fisher price index I_F (defined below as a geometric mean of the Laspeyres and Paasche indices) is exact for the linearly homogeneous quadratic aggregator function $f(x) = (x^T Ax)^{0,5}$, where A is a symmetric matrix of constants (see Diewert (1976)). The quadratic function above is an example of a *flexible functional form* (i.e. a function that provides a second order approximation to an arbitrary twice continuously differentiable function). Since I_F is exact for a flexible functional form, it is said to be a *superlative* index number (see Diewert (1976)). In the papers of Afriat (1972), Pollak (1971) and Samuelson-Swamy (1974) we can find other examples of exact index numbers and also superlative index numbers (see Diewert (1976), von der Lippe (2007)). Under all above remarks, an estimate of the commodity substitution bias B_{csub} can be given by (see White (1999))

$$B_{csub} \approx I_{La} - I_F. \tag{5}$$

In general, we can use any other superlative index $I_{\sup}$ to calculate the above mentioned CPI bias, namely then (see Hałka, Leszczyńska (2011))

$$B_{csub} \approx I_{La} - I_{\sup}. \tag{6}$$

In this paper we compare the CPI biases calculated by using some known, superlative price indices. In particular, we use the Fisher price index, the Törnqvist price index (I_T), the Walsh price index (I_W) and some original price index formula, defined in the next part of the paper (I_B). These indices are as follows (see von der Lippe (2007))

$$I_F = \sqrt{I_{La} I_{Pa}}, \tag{7}$$

$$I_W = \frac{\sum_{i=1}^{N} p_i^t \cdot \sqrt{q_i^s q_i^t}}{\sum_{i=1}^{N} p_i^s \cdot \sqrt{q_i^s q_i^t}}, \tag{8}$$

$$I_T = \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^s}\right)^{\overline{w}_i}, \tag{9}$$

where

$$w_i^s = \frac{p_i^s q_i^s}{\sum_{k=1}^{N} p_k^s q_k^s}, \quad w_i^t = \frac{p_i^t q_i^t}{\sum_{k=1}^{N} p_k^t q_k^t}, \quad \overline{w}_i = \frac{1}{2}(w_i^s + w_i^t). \tag{10}$$

3. The original price index formula

Białek (2012a) proposes the following price index

$$I_B = \sqrt{I_L \cdot I_U}, \tag{11}$$

where the *lower* (I_L) and *upper index* (I_U) we define as follows

$$I_L = \frac{\sum_{i=1}^{N} \min(q_i^s, q_i^t) p_i^t}{\sum_{i=1}^{N} \max(q_i^s, q_i^t) p_i^s}, \tag{12}$$

$$I_U = \frac{\sum_{i=1}^{N} \max(q_i^s, q_i^t) p_i^t}{\sum_{i=1}^{N} \min(q_i^s, q_i^t) p_i^s}. \tag{13}$$

In the above-mentioned paper it is proved that the index I_B^P satisfies price dimensionality, commensurability, identity, the mean value test, the time reversal test and linear homogeneity (see von der Lippe (2007)). Moreover, there are some interesting relations between this index and other formulas. For example, in the paper of Białek (2013) it is also proved that

$$\frac{I_F}{I_B} = \left[\frac{I_U}{I_L}\right]^{\frac{1}{2} - \frac{Log(I_U / I_F)}{Log(I_U / I_L)}}, \tag{14}$$

and thus

$$\sqrt{\frac{I_L}{I_U}} \leq \frac{I_F}{I_B} \leq \sqrt{\frac{I_U}{I_L}}. \tag{15}$$

It leads to the following conclusion

$$\forall i \in \{1,2,...,N\}\ q_i^s \approx q_i^t \ \Rightarrow\ I_L \approx I_U \ \Rightarrow\ I_F \approx I_B. \tag{16}$$

In the paper of von der Lippe (2012) it is proved that the Marshall-Edgeworth price index I_{ME} can be written as a weighted arithmetic mean of I_L and I_U, namely

$$I_{ME} = \frac{\sum_{i=1}^{N} p_i^s \max(q_i^s, q_i^t)}{\sum_{i=1}^{N} p_i^s \max(q_i^s, q_i^t) + \sum_{i=1}^{N} p_i^s \min(q_i^s, q_i^t)} I_L + \frac{\sum_{i=1}^{N} p_i^s \min(q_i^s, q_i^t)}{\sum_{i=1}^{N} p_i^s \max(q_i^s, q_i^t) + \sum_{i=1}^{N} p_i^s \min(q_i^s, q_i^t)} I_U \tag{17}$$

In fact, we can make a much more general observation – it can be proved (see Białek (2013)) that each of the above-mentioned indices (Fisher, Laspeyres, Paasche, Marhall-Edgeworth, Walsh formulas) have values between the *lower* and *upper index*. Thus, the formula I_B, as a geometric mean of I_L and I_U, seems to be well formed. In our simulation study (see Białek (2013)) it is shown, that the Fisher index and the Białek's price formula approximate each other. In the Section 4 we use the mentioned superlative price indices and the I_B index for calculating the CPI substitution bias in simulation studies.

4. Simulation study

Simulation 1

Let us take into consideration a group of $N = 50$ commodities, where random vectors of prices and quantities are as follows[1] (we present below only the first five commodities):

---

[1] The specification of prices and quantities does not mean that $p_i^t / p_i^s = a = const$ and $q_i^t / q_i^s = h(a) = const$, because random components of vectors P^t and Q^t are generated after the generating vectors P^s and Q^s. In other words, firstly we generate values of vectors of prices and quantities twice obtaining two pairs: (P^s, Q^s) and (P'^s, Q'^s) and after that we assume $P^t = aP'^s$ and $Q^t = h(a)Q'^s$.

$P^s = [U(400{,}700), U(1000{,}6000), U(3{,}9), U(3000{,}7000), U(100{,}500),...]'$,

$Q^s = [U(30000{,}70000), U(100{,}500), U(300{,}900), U(20000{,}50000), U(300{,}900),...]'$,

$P^t = a \cdot [U(400{,}700), U(1000{,}6000), U(3{,}9), U(3000{,}7000), U(100{,}500),...]'$,

$Q^t = h(a) \cdot [U(30000{,}70000), U(100{,}500), U(300{,}900), U(20000{,}50000), U(300{,}900),...]'$

where $U(m,n)$ denotes a random variable with the uniform distribution which has values in an $[m,n]$ interval and $h(a)$ is some positive function of the parameter $a > 0$. We consider three cases:

Case 1: $h(a) = 1/a$ which means that prices and quantities are negatively correlated;

Case 2: $h(a) = 1$ which means that prices and quantities are uncorrelated and consumption is on the constant level;

Case 3: $h(a) = a$ which means that prices and quantities are positively correlated.

We consider these three cases although only the first one is the most common in practice. However, sometimes consumers stock up on commodities although they observe the rise in prices and in this case prices and quantities are positively correlated. We generate values of these vectors in $n = 100000$ repetitions. We get the results[1] presented in Tab. 1, Tab.2 and Tab. 3 and on Fig. 1.

Case 1

Table 1. Basic characteristics of the discussed CPI bias measures for the given values of parameter a

Characteristics	$I_{La} - I_F$	$I_{La} - I_T$	$I_{La} - I_W$	$I_{La} - I_B$
	$a = 0.2$			
Mean	0.003191	0.002693	0.003537	0.003323
Standard deviation	0.003921	0.003502	0.004200	0.004192
Volatility coefficient	1.228871	1.300712	1.187441	1.261510
	$a = 0.5$			
Mean	0.001521	0.001672	0.001381	0.001404
Standard deviation	0.005902	0.007801	0.005902	0.005898
Volatility coefficient	3.878860	4.665669	4.273692	4.199560

[1] To read more about mean value estimation and the bias of this estimation see Żądło (2006), Gamrot (2007), Małecka (2011) or Papież, Śmiech (2013).

Table 1. Basic characteristics of the discussed CPI bias measures for the given values of parameter a (cont.)

Characteristics	$I_{La} - I_F$	$I_{La} - I_T$	$I_{La} - I_W$	$I_{La} - I_B$
	$a = 1$			
Mean	0.002513	0.001069	0.002877	0.002756
Standard deviation	0.01165	0.01108	0.01179	0.01173
Volatility coefficient	4.62112	10.35642	4.09681	4.25761
	$a = 1.5$			
Mean	0.001744	0.002202	0.001910	0.001775
Standard deviation	0.017100	0.021628	0.017262	0.017180
Volatility coefficient	9.802840	9.821980	9.034320	9.676200
	$a = 2$			
Mean	0.006229	0.003213	0.006427	0.006033
Standard deviation	0.023712	0.022446	0.023920	0.023774
Volatility coefficient	3.806440	6.984651	3.721341	3.940150

Source: Own calculations using Mathematica 6.0.

Case 2

Table 2. Basic characteristics of the discussed CPI bias measures for the given values of parameter a

Characteristics	$I_{La} - I_F$	$I_{La} - I_T$	$I_{La} - I_W$	$I_{La} - I_B$
	$a = 0.2$			
Mean	-0.002382	-0.001877	-0.002532	-0.002517
Standard deviation	0.003318	0.002939	0.003414	0.003405
Volatility coefficient	1.392530	1.565551	1.348520	1.353140
	$a = 0.5$			
Mean	-0.002344	-0.002338	-0.002498	-0.002472
Standard deviation	0.006125	0.006040	0.006212	0.006156
Volatility coefficient	2.612710	2.582451	2.486670	2.490291
	$a = 1.5$			
Mean	-0.006760	-0.007030	-0.005637	-0.005642
Standard deviation	0.018090	0.017812	0.017742	0.017716
Volatility coefficient	2.675760	2.533680	3.147060	3.139941
	$a = 2$			
Mean	-0.004522	-0.004402	-0.005210	-0.005148
Standard deviation	0.023198	0.022387	0.023437	0.023359
Volatility coefficient	5.12930	5.084760	4.497961	4.537181

Source: Own calculations using Mathematica 6.0.

Case 3

Table 3. Basic characteristics of the discussed CPI bias measures for the given values of parameter a

Characteristics	$I_{La} - I_F$	$I_{La} - I_T$	$I_{La} - I_W$	$I_{La} - I_B$
	$a = 0.2$			
Mean	-0.001148	-0.001988	-0.001118	-0.001148
Standard deviation	0.002542	0.002984	0.002537	0.002542
Volatility coefficient	2.213850	1.500073	2.269740	2.213850
	$a = 0.5$			
Mean	-0.001893	-0.002050	-0.002111	-0.001893
Standard deviation	0.005976	0.005851	0.0061144	0.005984
Volatility coefficient	3.155621	2.853650	2.896052	3.160200
	$a = 1.5$			
Mean	-0.002294	-0.003171	-0.0024332	-0.002294
Standard deviation	0.0029232	0.003653	0.003022	0.002903
Volatility coefficient	1.274201	1.151940	1.242360	1.265571
	$a = 2$			
Mean	-0.003097	-0.003630	-0.002794	-0.002916
Standard deviation	0.003777	0.004212	0.003534	0.003627
Volatility coefficient	1.219540	1.160181	1.264651	1.243970

Source: Own calculations using Mathematica 6.0.

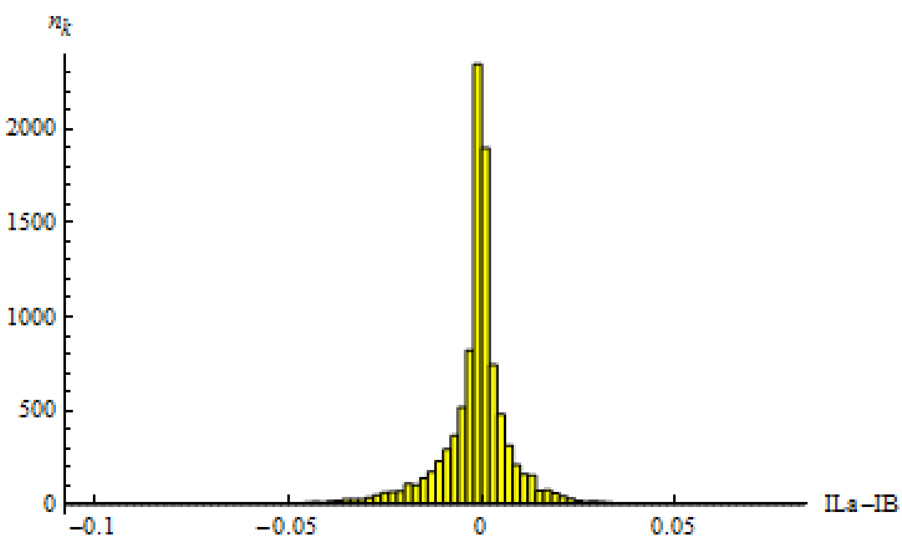


Figure 1. Histogram of $I_{La} - I_B$ in the case of $a = 1$

Source: Own calculations using Mathematica 6.0.

Simulation 2

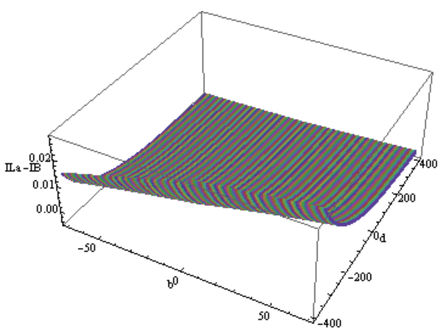
Let us take into consideration a group of only $N = 4$ commodities, where vectors of prices and quantities are as follows:

$P^s = [50,200,500,2500]'$, $P^t = [90,300,400,2000]'$,

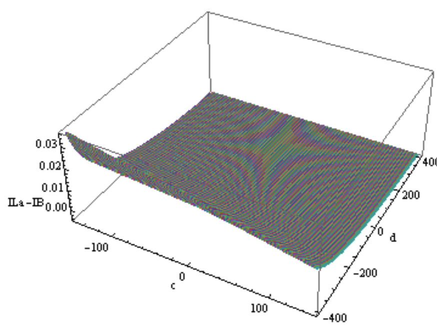
$Q^s = [300 + a, 90 + b, 200 + c, 500 + d]'$ and

$Q^t = [300,90,200,500]'$, for some real parameters a, b, c, d. The difference $I_{La} - I_B$ depending on these parameters is presented by Fig.2.

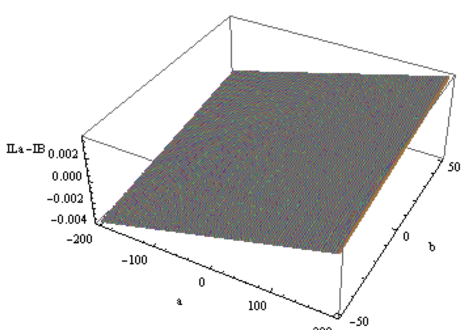
2.1. $c = d = 0$



2.2. $b = d = 0$



2.3. $b = c = 0$



2.4. $a = d = 0$

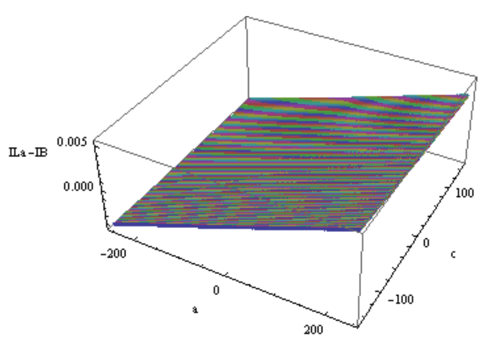
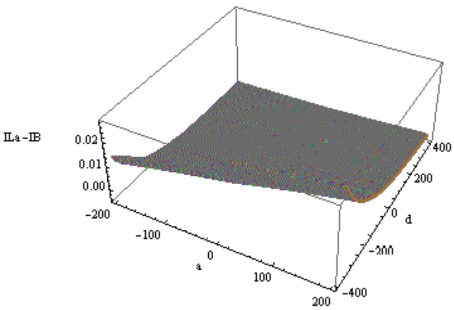


Figure 2. The CPI bias calculated as $I_{La} - I_B$ depending on parameters a, b, c and d

2.5. $a = c = 0$



2.6. $a = b = 0$

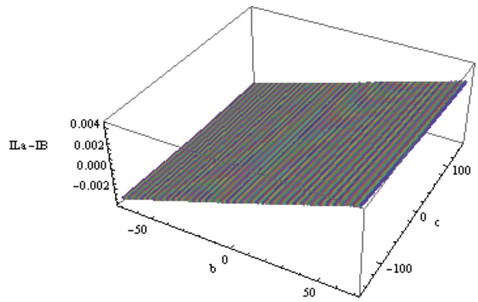


Figure 2. The CPI bias calculated as $I_{La} - I_B$ depending on parameters a, b, c and d (cont.)

Source: Own calculations using Mathematica 6.0.

5. Conclusions

In the simulation 1, case 1 (see Tab.1) we observe a positive expected value of the commodity substitution bias. This observation corresponds to the results coming from the report of the Boskin Commission, where we can find the estimated value of the commodity substitution bias on the level of 0,004 (see Boskin *et al.* (1996)). The similar conclusion can be also found in Cunnigham[1] (1996) or Crawford[2] (1998). We can notice that only in this case the estimator of the CPI bias calculated as $I_{La} - I_T$ is different in its expected value and has the highest volatility coefficient (the rest of estimators have similar values of this coefficient). Taking into consideration only the expected value of the commodity substitution bias we can find high similarity between the measures based on the Fisher and Białek formulas. Moreover, the scale of the commodity substitution bias does not seem to depend on the parameter a, which describes the changes in prices and quantities. In cases 2 and 3 (see Tab. 2 and Tab. 3 and Fig. 1) we observe negative expected value of the commodity substitution bias. Such situation can appear when the consumption does not depend on prices and is constant in time or quantities and prices are positively correlated. Then the real CPI is higher than the value obtained by Laspeyres formula (see researches by

[1] In this report the commodity substitution bias is in the interval 0 - 0.001.
[2] In this report the commodity substitution bias equals 0,001 and CPI is overestimated by 0.007.

Hałka, Leszczyńska[1] (2011), Ngasamiaku, Mkenda[2] (2009)). In the case 2 we can notice that bias measured by using the Walsh price index and the Białek index approximate one another. But the case 2 is not a real-life situation because it means that consumption remains on the constant level and is independent of changes in prices. It is worth noting that in case 3 we can also observe small differences between values of the CPI bias measures based on the Fisher and Białek formulas (as in the case 1). Taking into consideration good properties of the Białek formula (see also Białek (2012a), (2012b), (2013)) and the presented results we conclude that the I_B index can be a good alternative for the Fisher index in the CPI bias measurement.

In the simulation 2 we can notice that when parameters a, b, c and d increase then the Euklidean distance between quantities $d_e(Q^s, Q^t) = \sqrt{a^s + b^2 + c^2 + d^2}$ also increases and consequently the value of the difference $\left| I_{La}(Q^s, Q^t, P^s, P^t) - I_B(Q^s, Q^t, P^s, P^t) \right|$ becomes higher (see Fig. 2). For $a = b = c = d = 0$ we have $I_{La}(Q^s, Q^t, P^s, P^t) - I_B(Q^s, Q^t, P^s, P^t) = 0$.

Let us also notice (see Fig. 2) that the CPI substitution bias is an increasing function of parameters a and b but a decreasing function of parameters c and d. When the value of any parameter increases, we observe that the consumption decreases. Thus, the above observation confirms the conclusion from the simulation 1 if we notice that the first two products have higher prices and the last two products have lower prices at time t compared with time s.

Note

This article is based on the paper presented at the 7th Scientific Conference on Modelling and Forecasting of Socio-Economic Phenomena, May 7-10, 2013, Zakopane, Poland.

[1] In this paper the commodity substitution bias equals - 0.1 or 0.

[2] In this report the commodity substitution bias equals about - 0.0027.

REFERENCES

AFRIAT, S. N., (1972). The Theory of International Comparisons of Real Income and Prices, 13–69, In D. J. Daly (ed.), International Comparisons of Prices and Outputs, New York: Columbia University Press.

BIAŁEK, J., (2012a). Propozycja indeksu cen, Statistical News 7, 13–24, Central Statistical Office, Warsaw.

BIAŁEK, J., (2012b). Proposition of the general formula for price indices, Communications in Statistics: Theory and Methods, vol. 41:5, s. 943–952.

BIAŁEK, J., (2013). Simulation study of an original price index formula, Communications in Statistics: Simulation and Computation (in press, DOI: 10.1080/03610918.2012.700367).

BOSKIN, M. J., DULBERGER E. R., GORDON R. J., GRILICHES Z., JORGENSON D., (1996). Toward a More Accurate Measure of the Cost of Living, Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index.

CLEMENTS, K., IZAN H. Y., (1987). The Measurement of Inflation: A Stochastic Approach, Journal of Business and Economic Statistics, 5(3), 339–350.

DIEWERT, W. E., (1974). Functional forms for revenue and factor requirement functions, International Economic Review, 15, 119–130.

DIEWERT, W. E., (1976). Exact and superlative index numbers, Journal of Econometrics, 4, 114–145, North-Holland Publishing Company.

DIEWERT, W. E., (1983). The theory of the cost-of-living index and the measurement of welfare change, Price level measurement: proceedings of a conference sponsored by Statistics Canada, Eds. W. E. Diewert and C. Montmarquette, 163–233, Ottawa: Statistics Canada.

CRAWFORD, A., (1998). Measurement Biases in the Canadian CPI: An Update, Bank of Canada Review, Spring, 38–56.

CUNNINGHAM, A. W., (1996). Measurement biases in price indexes: an application to the UK's RPI, Bank of England, Working Paper Series 47.

DIEWERT, W. E., (1993). The economic theory of index numbers: a survey, Essays in index number theory, vol. 1, Eds. W. E. Diewert and A. O. Nakamura, 177–221, Amsterdam.

FRISCH, R., (1936). Annual survey of general economic theory: The problem of index numbers, Econometrica, 4, 1–39.

GRAMOT, W., (2007). Mean Value Estimation Using Two-Phase Samples with Missing Data in Both Phases, Acta Appl Math 96, 215–220.

HAŁKA, A., LESZCZYŃSKA, A., (2011). Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski, Gospodarka Narodowa 9, 51–75.

JORGENSON, D. W., SLESNICK, D. T., (1983). Individual and social cost of living indexes, Price level measurement: proceedings of a conference sponsored by Statistics Canada, Eds. W.E. Diewert and C. Montmarquette, 241–336, Ottawa: Statistics Canada.

KONYUS, A., BYUSHGENS, S., (1926). K probleme pokupatelnoi cili deneg, Voprosi Konyunkturi II: 1, 151–172.

MAŁECKA, M., (2011). Prognozowanie zmienności indeksów giełdowych przy wykorzystaniu modelu klasy GARCH, Ekonomista 6, 843–860.

NGASAMIAKU, A., MKENDA, W., (2009). An Analysis of Alternative Weighting System on the National Price Index in Tanzania: The Imlication to Poverty Analysis, Botswana Journal of Economics, 6, no. 10, 50–70.

PAPIEŻ, M., ŚMIECH, S., (2013). Causality in mean and variance within the international steam coal market, Energy Economics 36, 594–604.

POLLAK, R. A., (1971). The theory of the cost of living index, Research Discussion Paper no. 11, Office of Prices and Living Conditions, U. S. Bureau of Labor Statistics, Washington.

POLLAK, R. A., (1989). The theory of the cost-of-living index, Oxford: Oxford University Press.

SAMUELSON, P. A., SWAMY, S., (1974). Invariant economic index numbers and canonical duality: Survey and synthesis, American Economic Review 64, 566–593.

WHITE, A. G., (1999). Measurement Biases in Consumer Price Indexes, International Statistical Review, 67, 3, 301–325.

ŻĄDŁO, T. (2006). On prediction of total value in incompletely specified domains, Australian & New Zealand Journal of Statistics 48(3), 269–283.

VON DER LIPPE, P., (2007). Index Theory and Price Statistics, Peter Lang, Frankfurt, Germany.

VON DER LIPPE, P., (2012). Some short notes on a price index of Jacek Białek, Ekonometria (Econometrics), 35, 76–83.