UNSW Business School
Working Paper

Never Stand Still | Business School | Economics

UNSW Business School Research Paper No. 2015 ECON 15

Weekly versus Monthly Unit Value Price Indexes

W. Erwin Diewert
Kevin J. Fox
Jan de Haan

Weekly versus Monthly Unit Value Price Indexes

W. Erwin Diewert
University of British Columbia and UNSW

Kevin J. Fox*
UNSW

and

Jan de Haan
Statistics Netherlands and Delft University of Technology

17 July, 2015

Abstract

A previously overlooked source of potential bias in the Consumer Price Index (CPI) is described. We find that unit value (average) prices, commonly used for construction of the CPI should be constructed over the same period as the index to be constructed, rather than over an incomplete sub-period. The latter approach can lead to an upward bias in the CPI.

JEL Classification Numbers: C43, C82, E31.

Key Words: Elementary price indexes, aggregation, inflation.

*Contact author: School of Economics & Centre for Applied Economic Research, University of New South Wales, Sydney 2052 Australia, K.Fox@unsw.edu.au, Tel: +612 9385 3320.

Acknowledgements: Financial support from the Social Sciences and Humanities Research Council of Canada and the Australian Research Council's Discovery Program (DP150100830) is gratefully acknowledged.

1. Introduction

In constructing price indexes such as the Consumer Price Index (CPI), national statistical agencies make use of unit values at the elementary level of aggregation. A unit value is an average price over sales or purchases of a commodity that pertain to an economic unit (or group of units) and over a definite time period. Ratios of unit values, i.e. unit value indexes, "are frequently used by countries as surrogates for price changes at the elementary level of aggregation" (Silver, 2010, p. S210). The specification of a unit value involves accurately defining the commodity being aggregated, the economic agents that are involved in the aggregation and the time period over which transactions are aggregated. As scanner data from retail outlets and from households becomes more widely available, statistical agencies are either using or considering using unit value prices for narrowly defined commodities in place of their traditional sampling of prices in retail outlets approach.

The following paragraphs of the *Consumer Price Index Manual* (ILO and others, 2004) explain the importance of choosing the "right" unit value concept for the construction of a CPI:

"**20.15** It is now necessary to discuss a problem that arises when detailed data on individual transactions are available, either at the level of the individual household or at the level of an individual outlet. Recall that Chapter 15 introduces the price and quantity indices, $P(p^0,p^1,q^0,q^1)$ and $Q(p^0,p^1,q^0,q^1)$. These (bilateral) price and quantity indices decompose the value ratio V^1/V^0 into a price change part $P(p^0,p^1,q^0,q^1)$ and a quantity change part $Q(p^0,p^1,q^0,q^1)$. In this framework, it is taken for granted that the period t price and quantity for commodity i, p_i^t and q_i^t, respectively, are well defined. These definitions are not, however, straightforward since individual consumers may purchase the *same* item during period *t at different prices*. Similarly, if one considers the sales of a particular shop or outlet that sells to consumers, the same item may sell at very different prices during the course of the period. Hence before a traditional bilateral price index of the form $P(p^0,p^1,q^0,q^1)$ considered in previous chapters of this manual can be applied, a non-trivial time aggregation problem must be resolved in order to obtain the basic prices p_i^t and quantities q_i^t that are the components of the price vectors p^0 and p^1 and the quantity vectors q^0 and q^1."

"**20.16** Walsh[1] and Davies (1924) (1932) suggested a solution to this time aggregation problem: in their view, the appropriate quantity at this very first stage of aggregation is the *total quantity*

[1] Walsh explained his reasoning as follows: "Of all the prices reported of the same kind of article, the average to be drawn is the arithmetic; and the prices should be weighted according to the relative mass quantities that were sold at them." Correa Moylan Walsh (1901; 96). "Some nice questions arise as to whether only what is consumed in the country, or only what is produced in it, or both together are to be counted; and also there are difficulties as to the single price quotation that is to be given at each period to each commodity, since this, too, must be an average. Throughout the country during the period a commodity is not sold at one price, nor even at one wholesale price in its principal market. Various quantities of it are sold at different prices, and the full value is obtained by adding all the sums spent (at the same stage in its advance towards the consumer), and the average price is found by dividing the total sum

purchased of the narrowly defined item and the corresponding price is the value of purchases of this item divided by the total amount purchased, which is a *narrowly defined unit value*."

The *Consumer Price Index Manual* adopted this narrowly defined unit value concept as the solution to the time aggregation problem and hence as the appropriate concept for the prices and quantities that should enter a bilateral index number formula at the second stage of aggregation.[2]

Here we explore an issue which, to the best of our knowledge, has not received any previous consideration, yet could be a potential source of significant bias in the construction of the CPI in many countries. Specifically, in constructing a monthly price index we theoretically examine the implications of using only data from a subsample of the month, such as data for a particular week, as is common statistical agency practice. This is motivated by a recent empirical observation by Fox and Syed (2015) who, using scanner data from U.S. supermarkets, found that the monthly price indexes from constructing unit values from only particular weeks of the month typically lie above the indexes from constructing unit values across all weeks of the month.[3]

We find that the unit value should be constructed over the same period as the aggregate index, otherwise there is likely to be an upward bias in the index.

2. The Main Result

We develop a relationship between a weekly unit value price index for a narrowly defined commodity and the corresponding monthly unit value price index for the same commodity.[4] The relationship indicates that if the monthly unit value price index is the right target index, then a "representative" weekly price index is likely to have an upward bias.

(or the full value) by the total quantities." Correa Moylan Walsh (1921; 88). Fisher (1922; 318) also endorsed the use of unit values at the first stage of aggregation.

[2] For additional references to the ongoing research on unit values and index number theory, see Dalen (1992), Diewert (1995), Balk (1998), Reinsdorf (2009), Bradley (2005), de Haan (2004), Silver (2010) and Ivancic and Fox (2013).

[3] Fox and Syed (2015) used the IRI Academic Data Set (Bronnenberg, Kruger and Mela 2008) for the period 2001- 2011 for six major U.S. cities and ten broad product categories. Aggregating using the Jevons index at the elementary level and the Törnqvist index at higher levels, consistent with the approach of the U.S. Bureau of Labor Statistics in CPI construction, they found that by the end of the sample the upward bias caused by using only data from a particular week in constructing unit values could be more than 12 percentage points, or more than 1 percentage point per year on average; see Fox and Syed (2015), figure 3 .

[4] Our "monthly" index is defined as an aggregate of four weekly indexes. Since months are not precisely equal to the sum of four weeks, a true monthly unit value price index will not be precisely equal to our definition of a monthly index.

Let $p_w^t > 0$ denote the unit value price for a narrowly defined commodity in week w (w = 1,2,3,4) for month t where t = 0,1. Let $q_w^t > 0$ equal the total quantity sold of the commodity in week w of month t. For each choice of week w, we can define a *unit value price relative* r_w between months 0 and 1:

$$(1)\ r_w \equiv p_w^1/p_w^0 \equiv P_w; \qquad w = 1,\ldots,4,$$

where P_w is a month over month *weekly unit value price index*, which is identical to the unit value price relative r_w for the case of a single narrowly defined commodity.

The corresponding *monthly unit value price index* P_D for the same narrowly defined commodity can be defined using the weekly unit value prices and quantities as follows:

$$(2)\ P_D \equiv [\Sigma_{w=1}^4\, p_w^1 q_w^1/\Sigma_{w=1}^4\, q_w^1]/[\Sigma_{w=1}^4\, p_w^0 q_w^0/\Sigma_{w=1}^4\, q_w^0].$$

We labelled this index as P_D since Drobisch (1871) favoured this index number formula. Note that the Drobisch quantity index Q_D that matches up with the monthly unit value index defined by (2) is simply the monthly quantity ratio; i.e., we have:

$$(3)\ Q_D \equiv \Sigma_{w=1}^4\, q_w^1/\Sigma_{w=1}^4\, q_w^0.$$

Before we proceed to our analysis, it is necessary to develop an identity that was initially used by Bortkiewicz (1923; 374-375). First we define the monthly Laspeyres (1871) price and quantity indexes between months 0 and 1 that aggregate over the four weekly unit value prices and quantities. The ordinary *Laspeyres price index*, P_L, relating the prices of month 1 to those of month 0, can be defined as a share weighted average of the price relatives as follows:

$$(4)\ P_L \equiv \sum_{w=1}^4 p_w^1 q_w^0/\sum_{w=1}^4 p_w^0 q_w^0 = \sum_{w=1}^4 s_w^0\,(p_w^1/p_w^0) = \Sigma_{=1}^4\, s_w^0\, r_w \equiv r^*$$

where the *month 0 weekly expenditure shares* s_w^0 are defined as follows:

$$(5)\ s_w^0 \equiv p_w^0 q_w^0/\sum_{k=1}^4 p_w^0 q_w^0\ ; \qquad w = 1,\ldots,4.$$

Define the *nth quantity relative* t_n as the ratio of the quantity of the commodity sold in week w of month 1, q_w^t, to the corresponding quantity sold in week w of month 0, q_w^0, as follows:

$$(6)\ t_w \equiv q_w^1/q_w^0\ ; \qquad w = 1,\ldots,4.$$

The *Laspeyres quantity index*, Q_L, that compares weekly quantities in month 1 to the corresponding weekly quantities in month 0, using the prices of month 0, p^0, as weights can be defined as a share weighted average of the quantity ratios t_n as follows:

$$(7)\ Q_L \equiv \sum_{w=1}^4 p_w^0 q_w^1/\sum_{w=1}^4 p_w^0 q_w^0 = \sum_{w=1}^4 s_w^0\ (q_w^1/q_w^0) = \Sigma_{=1}^4\ s_w^0\ t_w \equiv t^*.$$

Using the above definitions, it is straightforward to establish the following *covariance identity*, which dates back to Bortkiewicz (1923) at least:

$$\begin{aligned}(8)\ \mathrm{Cov}(r,t,s^0) &\equiv \sum_{w=1}^4 s_w^0\ (r_w - r^*)(t_w - t^*)\\ &= \sum_{w=1}^4 s_w^0\ r_w t_w - t^* r^*\\ &= \sum_{w=1}^4 s_w^0\ r_w t_w - P_L Q_L.\end{aligned}$$

Although the sign and magnitude of the above weighted covariance between price change and quantity change can only be empirically determined, *purchaser substitution effects* will almost always lead to a negative covariance in empirical applications.[5]

Now we are ready to relate the monthly unit value price index defined by (2) for the narrowly defined commodity to the corresponding weekly unit value price indexes defined by (1). We start off with definition (2) which defines the monthly unit value price index:

$$\begin{aligned}(9)\ P_D &\equiv [\Sigma_{w=1}^4\ p_w^1 q_w^1/\Sigma_{w=1}^4\ q_w^1]/[\Sigma_{w=1}^4\ p_w^0 q_w^0/\Sigma_{w=1}^4\ q_w^0]\\ &= [\Sigma_{w=1}^4\ p_w^1 q_w^1/\Sigma_{w=1}^4\ p_w^0 q_w^0]/Q_D && \text{using definition (3)}\\ &= [\Sigma_{w=1}^4\ (p_w^1/p_w^0)(q_w^1/q_w^0)p_w^0 q_w^0/\Sigma_{n=1}^4\ p_n^0 q_n^0]/Q_D\\ &= [\Sigma_{w=1}^4\ s_w^0(p_w^1/p_w^0)(q_w^1/q_w^0)]/Q_D && \text{using definitions (5)}\\ &= [\mathrm{Cov}(r,t,s^0) + P_L Q_L]/Q_D && \text{using (1), (6) and (8)}\\ &= [\mathrm{Cov}(r,t,s^0)/Q_D] + P_L\,[Q_L/Q_D]\\ &= [\mathrm{Cov}(r,t,s^0)/Q_D] + P_w\,[P_L/P_w][Q_L/Q_D]\end{aligned}$$

[5] If the item is subject to periodic price discounting or is a seasonal item with large fluctuations in price, the covariance term can be fairly large.

$$= \alpha + \beta_w P_w \qquad \text{for } w = 1,2,3,4$$

for vectors $r = [r_1,\ldots, r_4]$, $t = [t_1,\ldots, t_4]$ and $s^0 = [s_1^0,\ldots, s_4^0]$, and where P_w is the weekly unit value price index defined in (1) and α and the β_w are defined as follows:

$$(10)\ \alpha \equiv \mathrm{Cov}(r,t,s^0)/Q_D\ ;$$
$$(11)\ \beta_w \equiv [P_L/P_w][Q_L/Q_D]\ ; \qquad w = 1,2,3,4.$$

As noted above, α will almost always be negative since the covariance $\mathrm{Cov}(r,t,s^0)$ is almost always negative and the Drobisch quantity index Q_D will be positive and typically close to one. The term β_w is equal to the product of two terms, $[P_L/P_w][Q_L/Q_D]$. The first term is equal to the ratio of the Laspeyres price index of the weekly unit value indexes, P_L, to the weekly unit value price index $P_w \equiv p_w^1/p_w^0$ that corresponds to week w. We know that a share weighted average $(\sum_{w=1}^4 s_w^0 (p_w^1/p_w^0) = \sum_{w=1}^4 s_w^0 P_w)$ of the weekly indexes P_w is in fact equal to P_L so that the harmonic mean of these P_L/P_w ratios, $\{\sum_{w=1}^4 s_w^0 [P_L/P_w]^{-1}\}^{-1}$, will equal one.[6] The second term, Q_L/Q_D, is the ratio of the monthly Laspeyres index that treats the unit values of each week as separate commodities to the Drobisch quantity index Q_D, and this will typically be close to one. Thus on average, the weekly unit value price indexes, P_w, will tend to be above the corresponding monthly unit value price index, P_D, with the magnitude of the differences depending primarily on the size of the covariance substitution effects.

3. Conclusion

The clear implication from these results is that statistical agencies should use monthly unit value prices in producing their monthly indexes rather than selecting a single week's (or similar sub-period's) unit value index. The latter indexes are likely to have an upward bias.[7] The implication is the same for statistical agencies that construct quarterly price indexes. The period over which the unit value is constructed should coincide with the frequency of the index.

---

[6] Since $P_L = \Sigma_{w=1}^4 s_w^0 P_w$, $1 = \Sigma_{w=1}^4 s_w^0 (P_w/P_L) = \Sigma_{w=1}^4 s_w^0 (P_w/P_L)^{-1}$ and thus $1 = [\Sigma_{w=1}^4 s_w^0 (P_w/P_L)^{-1}]^{-1}$ as well.

[7] Note that this has implications for statistical agencies that use incomplete periods of any length in constructing their unit values. This could be said of previous U.S. Bureau of Labor Statistics (BLS) practice, when data only from incomplete months were used: "Before 2004, data collection covered three pricing periods, each comprising 6 business days in most months and 5 days in November and December. Consequently, the last scheduled data collection was usually the 18th business day of the month. Beginning with data for January 2004, the three pricing periods now are of variable length and end on the last business day of the month." (BLS 2007, p. 17)

References

Balk, B. M. (1998), “On the Use of Unit Value Indices as Consumer Price Subindices”, Paper presented at the Fourth Meeting of the Ottawa Group at the International Conference on Price Indices, Washington, April 22–24.

Bortkiewicz, L. von (1923), “Zweck und Struktur einer Preisindexzahl”, *Nordisk Statistisk Tidsskrift* 2, 369-408.

Bradley, R. (2005), “Pitfalls of Using Unit Values as a Price Measure or Price Index”, *Journal of Economic and Social Measurement* 20, 39–61.

Bronnenberg, B.J., M.W. Kruger and C.F. Mela (2008), “Database Paper: The IRI Marketing Data Set”, *Marketing Science* 27, 745-748

Dalen, J., “Computing Elementary Aggregates in the Swedish Consumer Price Index,” *Journal of Official Statistics* 8, 129–47, 1992.

Davies, G.R. (1924), “The Problem of a Standard Index Number Formula”, *Journal of the American Statistical Association* 19, 180-188.

Davies, G.R. (1932), “Index Numbers in Mathematical Economics”, *Journal of the American Statistical Association* 27, 58-64.

Diewert, W.E. (1995), “Axiomatic and Economic Approaches to Elementary Price Indexes,” NBER Working Paper No. W5104.

Drobisch, M.W. (1871), “Über die Berechnung der Veränderungen der Waarenpreise und des Geldwerths”, *Jahrbücher für Nationalökonomie und Statistik* 16, 143–156.

Fisher, I. (1922), *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Fox, K.J. and I. Syed (2015), “Price Discounts and the Measurement of Inflation”, paper presented at the 14th Ottawa Group Meeting, Urayasu, Japan, 20-22 May.

de Haan, J. (2004), “Estimating Quality-Adjusted Unit Value Indices: Evidence from Scanner Data,” Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, June 30–July 3.

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

Ivancic, L. and K.J. Fox (2013), “Understanding Price Variation Across Stores and Supermarket Chains: Some implications for CPI aggregation methods”, *Review of Income and Wealth* 59, 629-647.

Laspeyres, E. (1871), “Die Berechnung einer mittleren Waarenpreissteigerung”, *Jahrbücher für Nationalökonomie und Statistik* 16, 296-314.

Reinsdorf, M. (1999), “Using Scanner Data to Construct CPI Basic Component Indexes”, *Journal of Business and Economic Statistics* 17, 152–60.

Silver, M. (2009), "An Index Formula Problem: The Aggregation of Broadly Comparable Items", IMF Working Paper 09/19.

Silver, M. (2010), "The Wrong and Rights of Unit Value Indices," *Review of Income and Wealth* 56, Special Issue 1, S206–23.

Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.

Walsh, C.M. (1921), *The Problem of Estimation,* London: P.S. King & Son.