The Effects of the Frequency and Implementation Lag of Basket Updates on the Canadian CPI

Ning Huang, Waruna Wimalaratne and Brent Pollard

Statistics Canada, Consumer Prices Division

14th Ottawa Group Meeting
Tokyo, Japan
20-22 May 2015

Table of Contents

List of Tables

List of Figures

The Effects of the Frequency and Implementation Lag of Basket Updates on the Canadian CPI

Ning Huang, Waruna Wimalaratne and Brent Pollard[1]

April 17, 2015

1. Introduction

The Consumer Price Index (CPI) is the most widely used indicator of price change in Canada. Its various uses include its function as a general indicator of inflation in Canada and as a tool for adjusting incomes, wages and other payments to ensure that purchasing power remains the same in an environment of changing prices. Serving a variety of purposes, it is of interest to governments, unions, business organizations, research institutions and the general public.

In line with the practices of other national statistical agencies, the Consumer Prices Division (CPD) at Statistics Canada uses the Lowe index formula for aggregating its CPI at the upper level. The Lowe index formula, often described as a "Laspeyres–type" formula, is a fixed–basket formula. This means that the quantity and quality of the goods and services included in the CPI basket must be unchanged or equivalent within the life span of a CPI basket.

As a good choice for the fixed–basket concept, the Lowe formula offers CPI compilers a simple and convenient way to compile composite price indices in a timely manner. Nevertheless, its inherent limitations must be taken into consideration. For example, it cannot account for consumers' product substitution, it experiences delay in reflecting the effects of new products on consumer price change, and it has difficulty in fully accounting for changes in the quality of consumer products. Due to these and other limitations, the official CPI, published by Statistics Canada, is not a true measure of actual changes in the cost of living.

The cost-of-living index (COLI), derived from the standpoint of economic theory, is based upon the assumption of consumers' utility optimization, which supposes that consumers will structure

[1] The authors thank Ross Beck-MacNeil, Xin Ha, Jean Le Moullec, Mathieu Lequain, Sue Morris, Philip Smith, Faouzi Tarkhani, Amanda Wright, Alice Xu and two internal reviewers at Statistics Canada, for their helpful discussions and comments. Special thanks to Erwin Diewert for his insightful comments. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of Statistics Canada.

their purchases to attain the optimal utility, or satisfaction, with the given prices of goods and services. It then measures the change in the minimum cost of maintaining this level of utility across two periods.

The difference between the official CPI and an underlying COLI[2], which can be approximated by a class of superlative indices, is called *measurement bias*[3]. The main types of measurement bias include commodity-substitution bias, outlet-substitution bias, quality-change bias and new-goods bias.

These types of bias arise from the fact that any basket weights held constant over more than one period do not necessarily reflect the types of purchases that consumers actually make to attain the same level of satisfaction when relative prices change. A fixed-basket index, therefore, normally fails to account for the changes in purchasing patterns in a timely manner, and measures only the average price movement based on a specifically defined basket, resulting in measurement bias. A COLI, on the other hand, accounts for changes in consumer purchasing patterns when measuring price movements over time.

Since the CPI is the most commonly used indicator for tracking overall price change in Canada, measurement bias in the CPI is an important issue to both its users and compilers. In addition, given the varying uses of the CPI, research on the measurement bias in the Canadian CPI is conducted regularly by some of its users.[4]

This paper focuses on the investigation of commodity-substitution bias. Generally speaking, without changing the formula for compiling the CPI, this type of bias could be reduced by updating the CPI basket more frequently and by implementing the basket in a timelier fashion. Both of these methods allow a more accurate reflection of the changes in purchasing patterns due to consumers' substitution between different combinations of goods and services. In the existing literature associated with commodity-substitution bias, there are only a limited number of studies examining the impact of the frequency and delay of implementing a new basket on the CPI. This is likely due to the difficulties associated with acquiring the data.

---

[2] According to the *Consumer Price Index (CPI) Manual (ILO et al., 2004),* a group of "superlative" price indices, such as Fisher, Walsh and Törnqvist index, are expected to provide "fairly close" approximations to the underlying Cost of Living Index. Thus, they are recommended by the CPI ILO Manual as the "target indices" for the upper-level index.
[3] In this study, only the measurement bias associated with the upper-level aggregation is discussed. Apart from this bias, there could also be sampling and other non-sampling bias in the estimated elementary indices and estimated basket weights. Measurement bias can be measured in terms of index level and index growth rate. In this paper, commodity-substitution bias is analysed and reported in both ways depending on the context.
[4] For instance, the most recent paper by the Bank of Canada is "Measurement Bias in the Canadian Consumer Price Index: An Update" by Patrick Sabourin, published by "Bank of Canada Review" in summer 2012.

It is, however, widely recognized that more frequent basket updating and faster implementation will lead to an index that approximates a superlative measure more closely. Généreux (1983) compared a chained Laspeyres series with eight basket updates against a chained Laspeyres series with only one basket update over the period from1957 to1978 using Canadian data. The comparison was made between the Laspeyres index and Fisher index. He concluded “what appears to be desirable is not necessarily a more frequent updating of the CPI basket but a more timely one”. [5] For example, implementing the new weights in the years they refer to could considerably reduce the commodity-substitution bias. A more recent study by Greenlees and Williams (2009) showed that annual basket updating generated an index that more closely resembled a target index,[6] when compared to less frequent updating. They also found that the advantages of using more timely baskets are not offset by any increase in index volatility or instability. Ho, Campion and Pike (2011) examined empirically the impact on the New Zealand CPI of reweighting at different frequencies and at different levels of the index structure using data from 2002 to 2008. They showed that frequent weight updates at sub-item level[7] and above generated CPI series that tracked the Fisher series most closely.

Since 2010, Statistics Canada has implemented a multi-stage program to advance the quality of the CPI, named the *CPI Enhancement Initiative*. As part of this initiative, effort has been put towards identifying and reducing the commodity-substitution bias. In 2013, a more frequent basket update schedule was implemented—from once every four years to once every two years. Additionally, the 2011 basket was introduced more quickly than past baskets—the time lag went from 16 months to 13 months. In this paper, we are interested in investigating the effect that changes such as these have on the quality of the CPI. Specifically, we aim to figure out whether we should accelerate the frequency of basket updates and reduce the implementation lag even further. Moreover, we want to know whether there exists an optimal month in which to introduce a new CPI basket.

The Canadian economy is a knowledge-based economy, which is therefore associated with dynamic technological changes. With the rapid applications of new technology and emergence of new products, consumers’ lifestyles and producers’ pricing strategies have also experienced significant changes. As a result, we expect that a CPI basket becomes outdated more rapidly in this type of economy than say an industrial economy. With this in mind, we question whether Généreux’s conclusions still hold and whether the empirical results from studies in other

[5] See Généreux 1983 (p.409).
[6] Greenlees and Williams (2009) calculated a chained Törnqvist index as a superlative target.
[7] The New Zealand CPI classification comprises groups, subgroups, classes, sections, subsections, items and subitems. There are nearly 700 subitems.

countries are also true for Canada. This study, therefore, also intends to identify the impact of different frequencies of basket updates and implementation lags on reducing upper-level commodity-substitution bias under the current economic situation, and to provide theoretical support for the potential future improvement upon the CPI at Statistics Canada.

The remainder of this paper is organized as follows: section two discusses the data sources and data construction methods; section three defines the target price index used in this study, which belongs to a group of superlative indices that closely approximate a COLI; section four addresses the effects of the frequency and implementation lag of basket updates on the Canadian CPI in detail; and section five concludes the paper.

2. Data Source and Construction

Since 1997, Statistics Canada has conducted the Survey of Household Spending (SHS) every year in the 10 provinces and every other year in the territories. This provides us with an opportunity to construct annual CPI baskets. Applying similar methodology and the same mapping rules[8] used in deriving the official CPI weights, we construct annual expenditure weights at the basic class level[9] for Canada over 12 years from 2000 to 2011. This data set allows us to investigate how the frequency of basket–updates and the lag with which the baskets are implemented affect the upper-level commodity-substitution bias.

2.1 Data Source

The two main elements required for the calculation of a price index series are prices and quantities. To this end, this study makes use of two main sources of data—the Consumer Price Index (CPI) and the Survey of Household Spending (SHS). The CPI provides data on the price indices for each of its measured goods and services at the basic class level of aggregation. The SHS data are used in constructing fixed–basket weights for the 12 years based on the 2005 CPI classification.

---

[8] The mapping rule is referring to the mapping or concordance of SHS product categories to CPI product categories.
[9] Basic classes are the lowest-level aggregates of products, chosen by Statistics Canada, for which a set of weights is fixed for the duration of the CPI basket.

2.1.1 Consumer Price Index

The "price" component of the index calculation comes from the CPI over the period from January 2000 to December 2013. The original price indices are unlinked price indices for each of the corresponding published CPI baskets. To facilitate the index reconstruction, the indices are linked together based on the classification of the 2005 basket and rebased to January 2000 = 100. The reconstructed indices therefore represent price movement from the price reference period of January 2000 to a given price observation month *t*.

2.1.2 Survey of Household Spending

The "quantity" component of the index comes from the SHS, which contains detailed information about consumer spending during a given reference year.[10] The survey is carried out in private households across the provinces on an annual basis.[11] The SHS sample has a cross-sectional design, and is selected from the Labour Force Survey sampling frame. The SHS is the main source of the expenditure weights data for the CPI.

2.2 Data Construction

2.2.1 Development of Basket Weights

In the first stage of this study, we constructed expenditure weights for Canada for the years without official CPI weights—2000, 2002 to 2004, 2006 to 2008 and 2010—using data from the SHS. Official CPI weight data were used whenever they were available, namely for the 2001, 2005, 2009, and 2011 baskets. However, some adjustments were made in order to align them with the 2005 classification of the CPI at the basic class level of aggregation. The 2005 classification was maintained across time to preserve uniformity and avoid complications arising from the introduction of new items.

The expenditure weights for basic classes are constructed by matching CPI basic classes with responses from the SHS in each year. However, complications arise in situations where the SHS

[10] Since 2010, the SHS has had an expanded diary component which includes the following: Food, parts of Transportation, Health and Personal Care, Recreation, and Alcohol and Tobacco. In the past, the diary component was generally available only for basket update years and was limited to food expenditures.
[11] The SHS became an annual survey starting from 1997. However, the SHS is conducted every two years in the Northern territories.

lacks data for a necessary response, or when there are changes in the method of aggregation used in the SHS between years.

Additionally, for non-official basket update years, some expenditure values are unavailable from the SHS. To impute these expenditure values, we utilise the backward and forward price updating method. Under this method, the missing expenditures are imputed by using a weighted average of expenditures with detailed SHS information. Relatively greater importance is assigned to expenditures in baskets from periods closer to the imputation period, while relatively smaller importance is assigned to expenditures of baskets from further periods.

Detailed expenditure data for food are generally available only for basket-update years[12]. Based on the backward and forward price updating procedure, the weights for intermittent years were constructed as follows:

$$
(1)\quad
\begin{aligned}
p_i^{2002}q_i^{2002} &\equiv (7/8)v_i^{2001}\frac{p_i^{2002}}{p_i^{2001}} + (1/8)v_i^{2009}\frac{p_i^{2002}}{p_i^{2009}}; && \text{i=1,...,N}\\
p_i^{2003}q_i^{2003} &\equiv (6/8)v_i^{2001}\frac{p_i^{2003}}{p_i^{2001}} + (2/8)v_i^{2009}\frac{p_i^{2003}}{p_i^{2009}}; && \text{i=1,...,N}\\
p_i^{2004}q_i^{2004} &\equiv (5/8)v_i^{2001}\frac{p_i^{2004}}{p_i^{2001}} + (3/8)v_i^{2009}\frac{p_i^{2004}}{p_i^{2009}}; && \text{i=1,...,N}\\
p_i^{2006}q_i^{2006} &\equiv (3/8)v_i^{2001}\frac{p_i^{2006}}{p_i^{2001}} + (5/8)v_i^{2009}\frac{p_i^{2006}}{p_i^{2009}}; && \text{i=1,...,N}\\
p_i^{2007}q_i^{2007} &\equiv (2/8)v_i^{2001}\frac{p_i^{2007}}{p_i^{2001}} + (6/8)v_i^{2009}\frac{p_i^{2007}}{p_i^{2009}}; && \text{i=1,...,N}\\
p_i^{2008}q_i^{2008} &\equiv (1/8)v_i^{2001}\frac{p_i^{2008}}{p_i^{2001}} + (7/8)v_i^{2009}\frac{p_i^{2008}}{p_i^{2009}}; && \text{i=1,...,N}
\end{aligned}
$$

where $v_i^t = p_i^t q_i^t$ is the estimated annual expenditure for category *i* in year *t* and $\frac{p_i^t}{p_i^0}$ is the annualized CPI for category *i* in year *t* compared with the price reference period 0. In this way, we implicitly assumed that:

[12] Detailed food data are available through a diary: survey respondents are given a diary for recording their food expenditures for two weeks with varying recall periods.

$$
(2) \quad
\begin{aligned}
q_i^{2002} &\equiv (7/8)\, q_i^{2001} + (1/8)\, q_i^{2009}; && i=1,\ldots,N \\
q_i^{2003} &\equiv (6/8)\, q_i^{2001} + (2/8)\, q_i^{2009}; && i=1,\ldots,N \\
q_i^{2004} &\equiv (5/8)\, q_i^{2001} + (3/8)\, q_i^{2009}; && i=1,\ldots,N \\
q_i^{2006} &\equiv (3/8)\, q_i^{2001} + (5/8)\, q_i^{2009}; && i=1,\ldots,N \\
q_i^{2007} &\equiv (2/8)\, q_i^{2001} + (6/8)\, q_i^{2009}; && i=1,\ldots,N \\
q_i^{2008} &\equiv (1/8)\, q_i^{2001} + (7/8)\, q_i^{2009}; && i=1,\ldots,N
\end{aligned}
$$

As mentioned earlier, because of the lack of detail data for food items, this procedure affected various aggregates and basic classes in all the years between official CPI basket updates. The following aggregates were affected:

- Food and non-alcoholic beverages purchased from stores
- Food purchased from restaurants

Similar strategies were employed for calculating the weight for the *mortgage interest cost* basic class as well as some components of clothing[13]. In the case of the mortgage interest cost index, where Statistics Canada uses a separate method of calculation, data are available only for the official basket reference years. As a remedy, weights for the remaining years were calculated using the same method as employed for food. For the *replacement cost* basic class, the SHS lacks detailed housing data for non-official basket-update years, thus a combination of internal and external data was used to calculate its value.[14]

Several product categories were added to the 2009 and 2011 baskets. To align them with the 2005 classification, the following basic classes were excluded from the 2009 to 2011 baskets:

- Multipurpose digital devices
- Telephone equipment

Furthermore, some sub-classes were excluded from their corresponding basic classes. These affected the following basic classes:

[13] Detailed “clothing and footwear” data were unavailable for some years. This affected basic classes 4101, 4102, 4103, 4201, 4202, 4203, and 4204 (see Table A.1 of the Appendix).
[14] Replacement cost = average value of home (excluding land) × number of households × depreciation rate. Detailed housing information, such as the average value of a home, is unavailable from the SHS in a non-basket update years. The average residential home price is taken from the Canadian Real Estate Association (CREA). The CREA value is then adjusted to exclude the value of land by using the Statistics Canada New Housing Price Index (NHPI). A depreciation rate of 1.5% is used for all calculations.

- Other household services
- Financial services

Table A.1 in the Appendix lists the classes which required special treatment.

2.2.2 Calculation of Price Indices

The second stage of this study involved calculating price indices for each basic class. Monthly price indices were calculated for each basic class with January 2000 =100 using existing data from the official CPI at the national level for the fourteen years from 2000 to 2013,

To preserve the 2005 classification, a number of strategies were applied in the absence of continuity in a series of price indices. Specific cases are documented as follows:

- Where the official CPI for a given time period contains less detailed information than in the 2005 classification, the price indices of the higher level aggregates were used in the place of all its subcomponents. For instance, in the case of basic classes denoted as *not elsewhere specified* (NES),[15] which are not included in earlier official CPI baskets and therefore have no associated price indices, the aggregated price indices were used. These basic classes are:
    - Household furnishings and equipment, NES
    - Other public transportation, NES
    - Other health care goods, NES
    - Cultural and recreational services, NES
    - Reading material and printed material, NES
    - Other alcoholic beverages purchased in stores, NES
- For the major aggregate "clothing", more detail was available in past baskets than in the 2005 basket. In this case, linking was done at the higher aggregates common to the 2005 basket.[16]

[15] The weights for these NES classes, which were not available in earlier basket updates, came from the Survey of Household Spending.
[16] The 1996 basket, which was used until Dec 2002, contained greater detail for clothing than following baskets. For example, *women's clothing* had six basic classes beneath it. However, with the 2001 basket, *women's clothing*

- Special adjustments were also made to the following indices to align them with the 2005 basket:
    - Other household services[17]
    - Financial services
    - Eye care goods
    - Eye care services

Table A.2 in the Appendix lists the classes which required special methods of calculation.

To verify how well the constructed data aligns with the official CPI data, we constructed a CPI in the same manner as the official published CPI. Only the 2001, 2005, 2009, and 2011 CPI baskets, updated at the same schedule as the published CPI, were used in this index calculation. The link months and implementing months for each basket are listed in Table 2.1:

Table 2.1: CPI weight reference years with implementation and link months

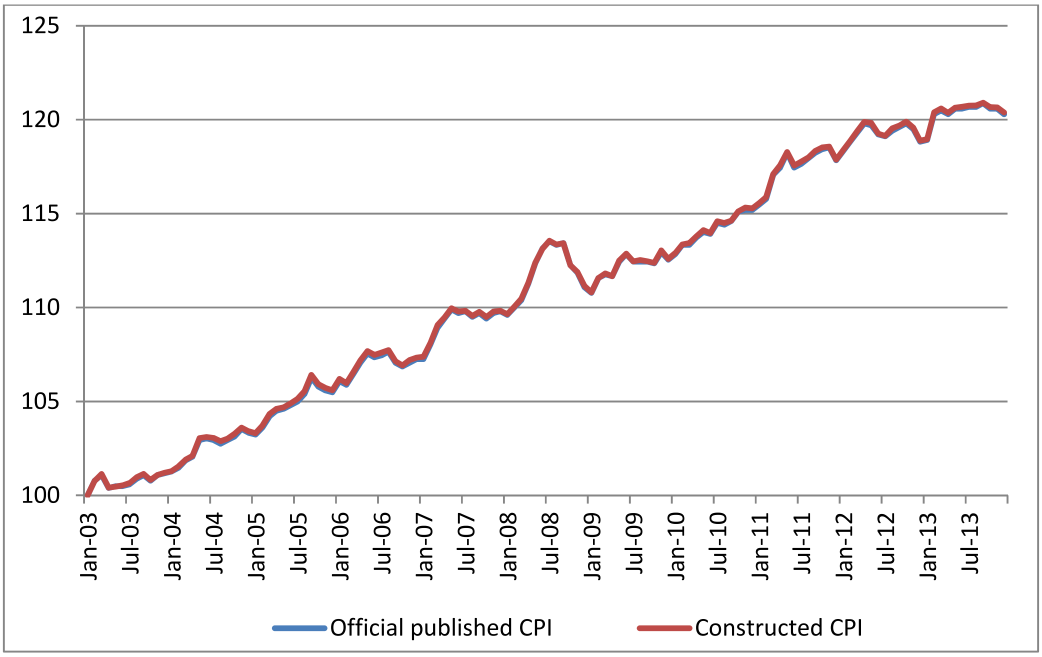
Weight reference year	2001	2005	2009	2011
Implementation month	Jan 2003	May 2007	May 2011	Feb 2013
Link month	Dec 2002	Apr 2007	Apr 2011	Jan 2013

Figure 2.1 below shows the comparison between the official published CPI and the CPI series compiled using the constructed data set with the official basket updating schedule.

Figure 2.1: Comparison of the official published CPI and the constructed CPI

was shown as a basic class only with no further breakdown. In this case, the higher level index (i.e. women's clothing) was used to link the 1996 basket to the 2001 basket.

[17] In the 2009 basket, the following items were added to the aggregate "other household services": legal services not related to dwelling services, funeral services, government services and retail club memberships. Since they were not comment to the 2005 basket, these sub-class indices were excluded from the calculation. In total, up to 26 indices out of 173 required special calculations, depending on the basket update year. See Appendix, Table A.2 for further details.



The two CPI series behave very similarly over the examined period. In fact, the constructed series lies nearly perfectly over the official series. In general, the index values calculated using the constructed data, which have been given special treatments in some basic classes, are very close to the official CPI values when rounded to one decimal point. This indicates that the constructed data set can represent the published CPI to a certain degree.

3. Target Index

To determine the magnitude of the substitution bias, first, it is necessary to select a target index with which to compare the estimates of this study.

The Fisher (3), Walsh (4) and Törnqvist (5) indices, defined as follows, belong to a small class of "superlative indices".

(3) Fisher: $$P_F^{t/0} = \left(P_L^{t/0} \times P_P^{t/0}\right)^{1/2} = \left(\frac{\sum_{i=1}^{N} p_i^t \times q_i^0}{\sum_{i=1}^{N} p_i^0 \times q_i^0} \times \frac{\sum_{i=1}^{N} p_i^t \times q_i^t}{\sum_{i=1}^{N} p_i^0 \times q_i^t} \right)^{1/2}$$

where $P_L^{t/0}$ is the Laspeyres index and $P_P^{t/0}$ is the Paasche index;

(4) Walsh: $P_W^{t/0} = \dfrac{\sum_{i=1}^{N} p_i^t \times \sqrt{q_i^t q_i^0}}{\sum_{i=1}^{N} p_i^0 \times \sqrt{q_i^t q_i^0}}$;

(5) Törnqvist: $P_T^{t/0} = \prod_{i=1}^{N} \left(\dfrac{p_i^t}{p_i^0} \right)^{\frac{1}{2}\left(s_i^0 + s_i^b\right)}$ where $s_i^0 = \dfrac{p_i^0 \times q_i^0}{\sum_{i=1}^{N} p_i^0 \times q_i^0}$ and $s_i^t = \dfrac{p_i^t \times q_i^t}{\sum_{i=1}^{N} p_i^t \times q_i^t}$

An important characteristic of superlative indices is that they treat the prices and quantities in both periods being compared symmetrically. They are therefore symmetrically weighted indices. Moreover, these three index number formulae are flexible and provide second-order approximation to each other. In other words, superlative indices tend to have similar properties, yield similar results and behave in very similar ways. In addition, they are expected to provide a close approximation to the underlying cost-of-living index (COLI)[18].

A COLI measures the change in the minimum cost of maintaining a given standard of living. It has some advantages over the Laspeyres-type index. For example, it, without delay, reflects substitution of goods and services cause by relative price change. It is, however, extremely difficult to compile for real time measurement of price change. As a close approximation to the unknown COLI, the superlative indices are recommended by the ILO CPI Manual as the theoretical target indices. The difference between the Laspeyres-type index, which does not permit commodity-substitution induced by relative price changes, and the target indices can be treated as a measure of commodity-substitution bias at the upper level of index aggregation.

In this study, we aim to compare chained-CPI series constructed by applying different baskets. The target indices are estimated by using the chain-linked Fisher, Walsh, or Törnqvist index number formulae with annual basket–updating, as detailed monthly expenditure data are unavailable. Using the Fisher index number formula $P_{ChF}^{2003+t/2003}$ as an example, we show how the chain-linked index between 2003 and 2011 is constructed:

[18] Diewert (1976) showed that superlative indices provide close approximations to any true cost of living price index if the underlying utility function is linear homogeneous.

(6)

$$P_{ChF}^{2003+t/2003} = \prod_{j=1}^{t} P_F^{2003+j/2003+j-1}$$

$$= \prod_{j=1}^{t} \left(\underbrace{\frac{\sum_{i=1}^{N} p_i^{2003+j} \times q_i^{2003+j-1}}{\sum_{i=1}^{N} p_i^{2003+j-1} \times q_i^{2003+j-1}}}_{P_L^{2003+j/2003+j-1}} \times \underbrace{\frac{\sum_{i=1}^{N} p_i^{2003+j} \times q_i^{2003+j}}{\sum_{i=1}^{N} p_i^{2003+j-1} \times q_i^{2003+j}}}_{P_P^{2003+j/2003+j-1}} \right)^{1/2} \qquad t=1, 2, ...8$$

where $P_{ChF}^{2003+t/2003}$ denotes the chained-Fisher index from 2003 to 2003+*t*; $P_F^{2003+j/2003+j-1}$ denotes the direct Fisher index from *2003+j-1* to *2003+j*; $P_L^{2003+j/2003+j-1}$ denotes the direct Laspeyres index from *2003+j-1* to *2003+j* and $P_P^{2003+j/2003+j-1}$ denotes the direct Paasche index from *2003+j-1* to *2003+j*; and *N* is the total number of goods and services included in the CPI basket. Note that $P_F^{2003/2003} = 100$. The chained-Walsh index and chained-Törnqvist index can be compiled similarly. Since the chained-superlative indices do not satisfy the transitivity test, certain degrees of chain drift exists in these indices.

The following table shows the empirical results from the calculation. Both index values and annual growth rates are listed in Table 3.1.

Table 3.1: Superlative Price Indices (2003=100)

Year (2003+*t*)	Fisher		Walsh		Törnqvist	
	Chained Index	Annual Inflation	Chained Index	Annual Inflation	Chained Index	Annual Inflation
2003	100.000		100.000		100.000	
2004	101.728	1.728	101.730	1.730	101.730	1.730
2005	103.746	1.984	103.750	1.986	103.750	1.986
2006	105.475	1.667	105.480	1.668	105.482	1.669
2007	107.401	1.826	107.409	1.829	107.410	1.828
2008	109.624	2.069	109.632	2.070	109.633	2.069
2009	109.670	0.042	109.684	0.047	109.688	0.050
2010	111.404	1.581	111.422	1.585	111.422	1.581
2011	114.389	2.679	114.408	2.680	114.405	2.677

Average Growth Rate[19] (from 2003 to 2011)	1.695		1.697		1.696

The results in Table 3.1 demonstrate that the three superlative indices behave in a similar way over the period from 2003 to 2011. In most years, the chained-Fisher price index is lower than the other two superlative indices. The empirical results also show that, in general, the differences between the chained-Walsh and chained-Törnqvist indices are smaller than those between the chained-Fisher index and these two indices, respectively.

In the next section, we will use Table 3.1 to estimate the upper-level substitution bias by comparing the target indices with the CPI series compiled using different CPI basket–updating schedules or different implementation lags. More specifically, the chained-Fisher index is used as an example to estimate the commodity-substitution bias in this paper.

4. Approaches to Reducing Commodity-Substitution Bias

In general, the commodity-substitution bias could be measured as the difference between the published CPI and the target index, which belongs to the class of "superlative indices", where the items in the baskets are held constant over time.[20]

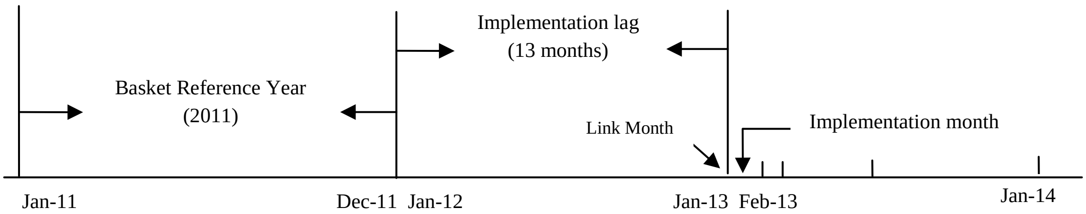
Two important sources[21] of the substitution bias could then be (i) the frequency with which the CPI basket is updated, and (ii) the time lag between the end of the basket reference year and implementation time of the basket (which is the period at which the basket is implemented in the CPI calculation). Using the 2011 basket update as an example, we illustrate the relationship among different time periods involved in the index calculation on the following timeline.

---

[19] The average growth rate is calculated as $\sqrt[8]{\left(P_{ChF}^{2011/2003}/100\right)}-1$.

[20] In this way, the estimated substitution bias would not be affected by the impact of changes in the specification and the appearance of the new products.

[21] The choice of index number formula would also affect the magnitude of the commodity-substitution bias in a country's CPI.



On the above timeline, the basket reference year (during which the SHS is conducted to collect the necessary information for the CPI basket) is 2011. January 2013 is the link month for introducing the 2011 CPI basket. The 2011 CPI basket is implemented with the February 2013 CPI, which is defined as the implementation month in this paper. The duration from January 2012 to January 2013 is the implementation lag, which, in this case, is 13 months.

In this section, we discuss in further detail how the frequency and implementation lag of the CPI basket affect the magnitude of the upper-level commodity-substitution bias.

4.1 Commodity-substitution Bias and the Frequency of Basket-update

4.1.1 Conceptual Framework to Measure the Impact of Basket-update Frequency

The CPI basket is designed to reflect consumers' spending patterns. As a result of both relative price changes and some long-term factors on consumers' spending behaviour, such as the demographic and technological changes, the basket might become out-of-date and less representative of current consumption patterns. The bias in a Lowe index is likely to increase as the basket weights age. Therefore, CPI baskets should be updated periodically to reflect the changes in these patterns.[22]

The frequency of basket–updating depends on the presumed magnitude of the commodity-substitution bias as well as on practical considerations, such as budgetary constraints and the feasibility of obtaining more recent data on consumer spending. In general, countries that are experiencing significant economic changes or rapid changes in consumption patterns should update their weights more frequently. In practice, this frequency varies from country to country. For example the UK and France update their baskets every year, the US updates its basket every

[22] The 2003 ICLS (International Conferences of Labour Statisticians) resolutions propose to update the CPI weights at least every 5 years.

two years, and New Zealand updates their CPI basket every three years. Statistics Canada started to update the basket biennially[23] as of 2013.

To identify the pure impact of the frequency of basket updating on the magnitude of the CPI bias, we fix the implementation lag at 13 months and vary only the frequency of basket–updating when calculating the All-items CPI[24] for the period from January 2002 to December 2013. For example, suppose *k* indicates how many times a CPI basket has been updated up to the current price observation period *t*. Then, with a basket–update frequency of once every two years, the $2000+2k$ basket is used in the CPI calculation as of February $2002+2k$ with January $2002+2k$ as the link month.

A direct Lowe index $P_{Lo}(p^0, p^t, q^b)$ can be defined in terms of a quantity vector $q^b \equiv \left[q^b_{1,} \dots q^b_{N,} \right]$, a base-period price vector $p^0 \equiv \left[p^0_{1,} \dots p^0_{N,} \right]$, and a current-period price vector $p^t \equiv \left[p^t_{1,} \dots p^t_{N,} \right]$:

(7) $$P_{Lo}(p^0, p^t, q^b) = \frac{\sum_{i=1}^{N} p_i^t \times q_i^b}{\sum_{i=1}^{N} p_i^0 \times q_i^b}$$

where *N* is the total number of goods and services included in the CPI basket.

It can also be written in terms of the hybrid expenditure shares as follows:

(8) $$P_{Lo}(p^0, p^t, s^{0:b}) = \frac{\sum_{i=1}^{N} p_i^t \times q_i^b}{\sum_{i=1}^{N} p_i^0 \times q_i^b} = \sum_{i=1}^{N} \left(\frac{p_i^t}{p_i^0} \right) \times \frac{p_i^0 \times q_i^b}{\sum_{i=1}^{N} p_i^0 \times q_i^b} = \sum_{i=1}^{N} \left(\frac{p_i^t}{p_i^0} \right) \times s_i^{0:b}$$

where the hybrid expenditure shares $s_i^{0:b}$ corresponding to the quantity weights vector q^b measured at base-period price vector p^0 are defined as:

(9) $$s_i^{0:b} = \frac{p_i^0 \times q_i^b}{\sum_{i=1}^{N} p_i^0 \times q_i^b}, \qquad i = 1, 2, \dots, N$$

[23] Statistics Canada previously updated the CPI basket every four years.
[24] The "All-items CPI" measures price changes of all the goods and services included in the Canadian CPI classification.

With the assumption that only the 2000 basket is used for the period from January 2002 to December 2013, the corresponding CPI, $P_{Lo}(p^0, p^t, q^{2000})$, can be compiled using the following equation based on equation (7):

$$(10) \qquad P_{Lo}(p^0, p^t, q^{2000}) = \frac{\sum_{i=1}^{N} p_i^t \times q_i^{2000}}{\sum_{i=1}^{N} p_i^0 \times q_i^{2000}}$$

Thus, the price index using only the 2000 CPI basket is simply calculated as the ratio of the cost of purchasing the same basket q^{2000} at different price vectors: p^0 and p^t.

If more than one basket, say basket b1 and b2, are in use, it is necessary to calculate the chain-linked Lowe index, where the indices calculated using different CPI baskets are linked together. To explain this concept, let $p^{y,m}$ be the elementary price vector for year y and month m; the chain-linked Lowe index for year $y \geq 2002$ and month m=1,2,…12, with every x years as the frequency of basket updating, is denoted as $P_{ChLo\_x}(y,m)$. The calculation of the chain-linked Lowe index depends on which basket is currently used and in which month it is linked to the previous basket. In general, a chain-linked Lowe index can be defined as:

$$(11) \qquad P_{ChLo\_x}(y,m) = P_{ChLo\_x}(link\_month) \times P_{Lo}(p^{link\_month}, p^{y,m}, q^b)$$

where $P_{ChLo\_x}(link\_month)$ is a chain-linked Lowe index for the link month that chains together indices using the current basket q^b and the previous baskets.

If the CPI basket is assumed to be updated every x years, where x can be 1, 2, 3, 4 or 5, after the adoption of the 2000 basket, equation (11) can be applied to compile the CPI series. With the implementation lag set to equal to 13 months, a new $2000+kx$ basket is introduced in February of year $2002+kx$ (with January (m=1) of year $2002+kx$ as the link month).[25] With these assumptions, the chain-linked Lowe index can be calculated by substituting the corresponding values in equation (11). This yields the following results:

$$(12) \qquad P_{ChLo\_x}(y,m) = P_{ChLo\_x}(2002+kx,1) \times P_{Lo}(p^{2002+kx,1}, p^{y,m}, q^{2000+kx})$$

[25] k denotes the number of times that the CPI basket updates between the price reference period *0* and current price observation period t such that $2002+kx \leq y$ (y is the year of the price index).

The current month (*y, m*) will determine the real values in equation (12). The first component of the right-hand side of equation (12), $P_{ChLo\_x}(2002+kx,1)$, is the link factor, which is also a chain-linked Lowe index for January of year *2002+kx*, the link month of the current basket (*2000+kx* basket); the second component, $P_{Lo}(p^{2002+kx,1}, p^{y,m}, q^{2000+kx})$, is the direct Lowe index comparing the current month (*y, m*) with the link month (*2002+kx, 1*), which is January of year $2002+kx$. The link factor can be defined as the product of several direct Lowe indices as follows:

(13)
$$P_{ChLo\_x}(2002+kx,1) = P_{Lo}(p^{0}, p^{2002+x,1}, q^{2000}) \times P_{Lo}(p^{2002+x,1}, p^{2002+2x,1}, q^{2000+x}) \times ... \times P_{Lo}(p^{2002+(k-1)x,1}, p^{2002+kx,1}, q^{2000+(k-1)x})$$

The price reference period *0* is assumed to be a period within the life span of the 2000 basket, for example January 2002.

We now describe how the chain-linked Lowe index can be constructed if the weights are updated every two years, that is *x*=2. Denote the chain-linked Lowe index for year *y* and month *m* with an update frequency of every two years, by $P_{ChLo\_2}(y,m)$. In this case, the direct Lowe index, which uses the 2000 basket only, is employed from February[26] 2002 to January 2004 with January 2002 as the link month. Thus, the chain-linked Lowe index defined by (14) below is, for the first 24 months running from February 2002 to January 2004, equal to the direct Lowe index:

(14)
$$P_{ChLo\_2}(y,m) = P_{Lo}(p^{2002,1}, p^{y,m}, q^{2000}) \quad \text{(with } y\text{=2002, 2003; and } m\text{=1,2,...12 and } y=2004; m=1)$$

The same direct Lowe index on the right-hand side of equation (14) is, therefore, used to define the chain-linked Lowe index for January 2004:

(15)
$$P_{ChLo\_2}(2004,1) = P_{Lo}(p^{2002,1}, p^{2004,1}, q^{2000})$$

This chain-linked Lowe index for January 2004 corresponds to the link factor that chains together indices using the 2000 basket and the 2002 basket. For the remaining months in 2004

---

[26] The 2000 basket is also applied to January of year 2002. It is the overlapping period that links the old and new CPI series. Applying (14), we have $P_{ChLo\_2}(2002,1) = P_{Lo}(p^{2002,1}, p^{2002,1}, q^{2000}) = 1$.

and 2005, the vector of annual quantity weights q^{2002} becomes available and the chain-linked Lowe index is defined as follows:

(16) $$P_{ChLo\_2}(y,m) = P_{ChLo\_2}(2004,1) \times P_{Lo}(p^{2004,1}, p^{y,m}, q^{2002}) \quad \text{(with } y\text{=2004, 2005; } m\text{=1, 2, ...,12; } y = 2006;\ and\ m = 1)$$

The chain-linked Lowe index for January 2006 is, therefore, defined as follows:

(17) $$P_{ChLo\_2}(2006,1) = P_{ChLo\_2}(2004,1) \times P_{Lo}(p^{2004,1}, p^{2006,1}, q^{2002})$$

The chain-linked Lowe index for January 2006 is the link factor that chains indices based on the 2004, 2002 and 2000 baskets. From February 2006 to January 2008, the vector of annual quantity weights q^{2004} becomes available and the chain-linked Lowe for this time span is defined as follows:

(18) $$P_{ChLo\_2}(y,m) = P_{ChLo\_2}(2006,1) \times P_{Lo}(p^{2006,1}, p^{y,m}, q^{2004}) \quad \text{(with } y = 2006,\ 2007;\ m = 1,\ 2, ..., 12;\ y = 2008;\ m = 1)$$

Again, the link factor chaining the indices together across baskets is the chain-linked Lowe index for January 2008 which continues to be defined by the right-hand side of equation (18) as follows:

(19) $$P_{ChLo\_2}(2008,1) = P_{ChLo\_2}(2006,1) \times P_{Lo}(p^{2006,1}, p^{2008,1}, q^{2004})$$

Continuing the above process, we can construct the chain-linked Lowe index for other months in the other years.

To show how the defined process works, the construction of a chain-linked Lowe index for a particular month, say August 2011, is provided as an example. Assume the basket–updating frequency is two (*x*=2) and the implementation lag is 13 months. The chained-Lowe index is then denoted by $P_{ChLo\_2}(2011,8)$. Based on the described process, the current period—August 2011—

is in the time span going from February 2010 to January 2012[27] and the associated vector of quantity weights is q^{2008} (with January 2010 as the link month). Henceforth, the chain-linked Lowe index $P_{ChLo\_2}(2011,8)$ can be constructed as:

$$(20)\qquad P_{ChLo\_2}(2011,8) = P_{ChLo\_2}(2010,1) \times P_{Lo}(p^{2010,1}, p^{2011,8}, q^{2008})$$

where $P_{ChLo\_2}(2010,1)$ is the link factor that chains together the price indices using the 2008 basket and the previous baskets. Based on equation (13), it can be defined as a product of the direct Lowe indices as follows:

$$(21)\qquad \begin{aligned} P_{ChLo\_2}(2010,1) = {} & P_{Lo}(p^{2002,1}, p^{2004,1}, q^{2000}) \times P_{Lo}(p^{2004,1}, p^{2006,1}, q^{2002}) \times \\ & P_{Lo}(p^{2006,01}, p^{2008,1}, q^{2004}) \times P_{Lo}(p^{2008,1}, p^{2010,1}, q^{2006}) \end{aligned}$$

The direct Lowe index on the right-hand side of (20) can be compiled based on equation (7) as follows:

$$(22)\qquad P_{Lo}(p^{2010,1}, p^{2011,8}, q^{2008}) = \frac{\sum_i p_i^{2011,8} \times q_i^{2008}}{\sum_i p_i^{2010,1} \times q_i^{2008}}$$

Next, we apply the same process described by equations (14) to (19) to another basket-updating frequency. This time, the chain-linked Lowe index for the same month—August 2011—but with a different basket–updating frequency, *x*=3, denoted by $P_{ChLo\_3}(2011,8)$, is considered. In this scenario, the annual basket q^{2000} is used for the first 36 months running from February 2002 to January 2005. Weights vector q^{2003} becomes available from February 2005 and is effective up to

[27] Based on the equation $2002 + kx = y$, the basket update periods can be derived by using ~~formula~~formula $k = \text{int}\left[(2011-2002)/x\right]$ (which takes the integer part of the calculated results) and the weight vector can be obtained from year *2000+kx*.

January 2008. Continuing with this process, we know that the chain-linked index $P_{ChLo\_3}(2011,8)$ is constructed by using weight vector q^{2009} as follows:

(23) $$P_{ChLo\_3}(2011,8) = P_{ChLo\_3}(2011,1) \times P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2009})$$

The chain-linked Lowe index $P_{ChLo\_3}(2011,1)$, which is the link factor chaining together the price indices using the 2009 basket and the previous baskets, can be derived from the following equation:

(24) $$\begin{aligned} P_{ChLo\_3}(2011,1) &= P_{ChLo\_3}(2008,1) \times P_{Lo}(p^{2008,1}, p^{2011,1}, q^{2006}) \\ &= P_{Lo}(p^{2002,1}, p^{2005,1}, q^{2000}) \times P_{Lo}(p^{2005,1}, p^{2008,1}, q^{2003}) \times P_{Lo}(p^{2008,1}, p^{2011,1}, q^{2006}) \end{aligned}$$

The direct Lowe index on the right-hand side of (23) is calculated as follows:

(25) $$P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2009}) = \frac{\sum_i p_i^{2011,8} \times q_i^{2009}}{\sum_i p_i^{2011,1} \times q_i^{2009}}$$

With the two CPI index values associated with different frequencies of basket-updating, the commodity-substitution bias in the different CPI index values can be then estimated by comparing the chain-linked Lowe index with the same target index. For example, let $Bias_{ChLo\_2}(2011,8)$ denote the commodity-substitution bias of the chain-linked Lowe index for August, 2011 with a basket updating frequency of every two years. Similarly, let $Bias_{ChLo\_3}(2011,8)$ denote the bias with a basket update frequency of every three years. The biases can be defined, respectively, as follows:

(26) $$Bias_{ChLo\_2}(2011,8) = P_{ChLo\_2}(2011,8) - P_{\text{Target}}(2011,8)$$

(27) $$Bias_{ChLo\_3}(2011,8) = P_{ChLo\_3}(2011,8) - P_{\text{Target}}(2011,8)$$

To compare the magnitude of the bias generated by different basket-updating frequencies, the following procedure is conducted:

(28)
$$
\begin{aligned}
& Bias_{ChLo\_2}(2011,8) - Bias_{ChLo\_3}(2011,8) \\
&= \left[P_{ChLo\_2}(2011,8) - P_{Target}(2011,8)\right] - \left[P_{ChLo\_3}(2011,8) - P_{Target}(2011,8)\right] \\
&= P_{ChLo\_2}(2011,8) - P_{ChLo\_3}(2011,8) \\
&= \left[P_{ChLo\_2}(2010,1) \times P_{Lo}(p^{2010,1}, p^{2011,8}, q^{2008})\right] - \left[P_{ChLo\_3}(2011,1) \times P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2009})\right] \\
&= P_{Lo}(p^{2002,1}, p^{2004,1}, q^{2000}) \times P_{Lo}(p^{2008,1}, p^{2010,1}, q^{2006}) \times \\
&\left\{ \begin{aligned}
&\left[\begin{aligned} &P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2002}) \times P_{Lo} \times (p^{2005,1}, p^{2006,1}, q^{2002}) \times P_{Lo}(p^{2006,1}, p^{2008,1}, q^{2004}) \times \\ &P_{Lo} \times (p^{2010,1}, p^{2011,1}, q^{2008}) \times P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2008}) \end{aligned}\right] - \\
&\left[\begin{aligned} &P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2000}) \times P_{Lo} \times (p^{2005,1}, p^{2006,1}, q^{2003}) \times P_{Lo}(p^{2006,1}, p^{2008,1}, q^{2003}) \times \\ &P_{Lo} \times (p^{2010,1}, p^{2011,1}, q^{2006}) \times P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2009}) \end{aligned}\right]
\end{aligned} \right\}
\end{aligned}
$$

To facilitate the comparison, all the direct Lowe indices in equation (28) are written in terms of the indices with the same price comparison periods.

From the right-hand side of equation (28), it can be seen that the two pairs of Lowe indices, $P_{Lo}(p^{2002,1}, p^{2004,1}, q^{2000})$ and $P_{Lo}(p^{2008,1}, p^{2010,1}, q^{2006})$], are identical; whereas, the other five pairs of Lowe indices measure the price movement over the same periods but use different quantity weight vectors:

- In three pairs of Lowe indices representing four years' of price change—from January 2004 to January 2005, $P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2002})$ and $P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2000})$; from January 2006 to January 2008, $P_{Lo}(p^{2006,1}, p^{2008,1}, q^{2004})$ and $P_{Lo}(p^{2006,1}, p^{2008,1}, q^{2003})$; and from January 2010 to January 2011, $P_{Lo} \times (p^{2010,1}, p^{2011,1}, q^{2008})$ and $P_{Lo} \times (p^{2010,1}, p^{2011,1}, q^{2006})$—those with a more frequent basket-updating schedule (x=2) use relatively more up-to-date quantity weight vectors;

- Whereas, of the other two pairs of indices corresponding to less than two years of price change—one from January 2005 to January 2006, $P_{Lo} \times (p^{2005,1}, p^{2006,1}, q^{2002})$ and $P_{Lo} \times (p^{2005,1}, p^{2006,1}, q^{2003})$, and the other from January 2011 to August 2011, $P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2008})$ and $P_{Lo}(p^{2011,1}, p^{2011,8}, q^{2009})$—those with less frequent basket–updates (x=3) use more up-to-date quantity weight vectors.

This simple comparison indicates that the chain-linked series with more frequent basket updates applies up-to-date quantity weights more often than the one with less frequent basket updates. Generally speaking, the price index compiled using a more outdated basket tends to *exceed* that with more up-to-date baskets due to the price-induced commodity-substitution. Thus, through this rough comparison, it is intuitively believed that more frequent basket–updating would generate lower commodity-substitution bias in general.

To identify conditions under which more frequent basket–updating would generate lower commodity-substitution bias, one of the pairs of the Lowe indices in equation (28) is compared:

$$
\begin{aligned}
&P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2002}) - P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2000}) \\
&= \frac{\sum_i p_i^{2005,1} \times q_i^{2002}}{\sum_i p_i^{2004,1} \times q_i^{2002}} - \frac{\sum_i p_i^{2005,1} \times q_i^{2000}}{\sum_i p_i^{2004,1} \times q_i^{2000}} \\
(29) \quad &= \sum_i \left(\frac{p_i^{2005,1}}{p_i^{2004,1}} \right) \times \left(\frac{p_i^{2004,1} \times q_i^{2002}}{\sum_i p_i^{2004,1} \times q_i^{2002}} - \frac{p_i^{2004,1} \times q_i^{2000}}{\sum_i p_i^{2004,1} \times q_i^{2000}} \right) \\
&= \frac{\sum_i \left(\frac{p_i^{2005,1}}{p_i^{2004,1}} - P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2002}) \right) \left(\frac{q_i^{2002}}{q_i^{2000}} - Q_{Lo}(p^{2004,1}, q^{2000}, q^{2002}) \right) \times s_i^{2004,1:2000}}{Q_{Lo}(p^{2004,1}, q^{2000}, q^{2002})}
\end{aligned}
$$

where the Lowe quantity index, $Q_{Lo}(p^{2004,1}, q^{2000}, q^{2002})$, is defined as:

(30) $$Q_{Lo}(p^{2004,1}, q^{00}, q^{02}) = \frac{\sum_i p_i^{2004,1} \times q_i^{2002}}{\sum_i p_i^{2004,1} \times q_i^{2000}},$$

and the hybrid expenditure shares $s_i^{2004,1:2000}$ are defined in terms of the year 2000 quantity vector q^{2000} at January 2004 prices:

(31) $$s_i^{2004,1:2000} = \frac{p_i^{2004,1} \times q_i^{2000}}{\sum_i p_i^{2004,1} \times q_i^{2000}}$$

The last line of equation (29) illustrates that the price relatives and quantity relatives are for two different periods. Provided that the price and quantity changes were for the same period (from 2000 to 2002), the right-hand side of equation (29) would be regarded as the covariance between the price relatives and the corresponding quantity relatives. If this covariance is negative (which is the usual case in the consumer context) and the price trend from 2000 to 2002, on average, is in the same direction as those going from January 2004 to January 2005, the Lowe index using up-to-date basket, $P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2002})$,will be lower than that using the out-dated basket, $P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2000})$.

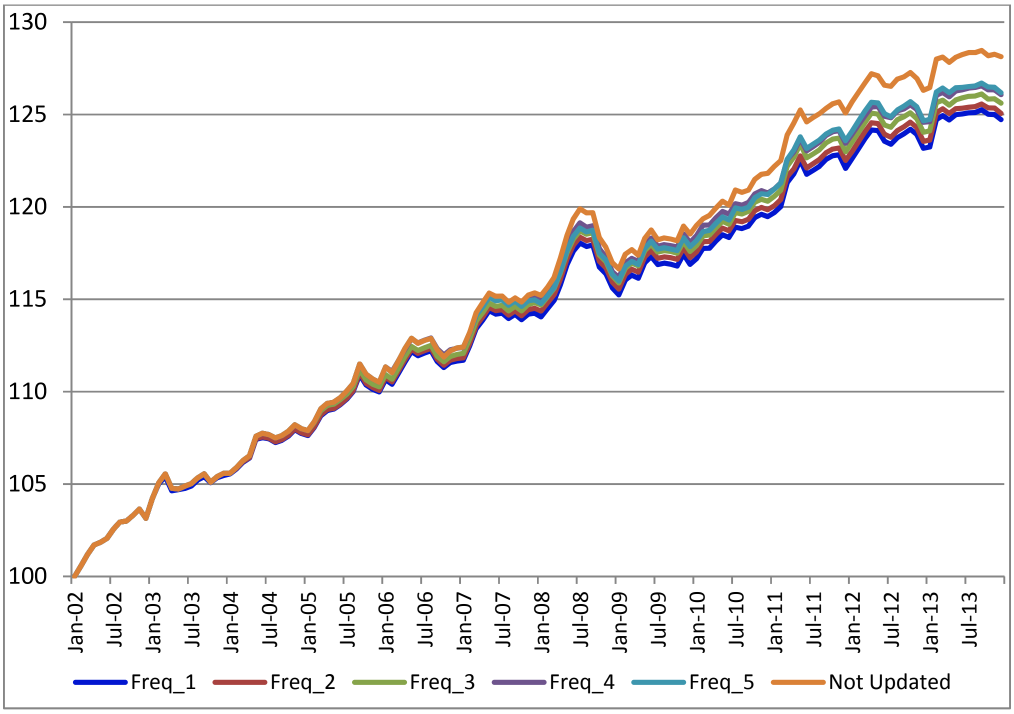
In short, the relationship between $P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2002})$ and $P_{Lo}(p^{2004,1}, p^{2005,1}, q^{2000})$ depends upon the persistent tendency of price change and the associated change in the consumers' expenditure patterns. This conclusion will also be true for the comparison of the other pairs of Lowe indices in equation (28). However, the determination of the sign of equation (28), which provides implications on the comparison of commodity-substitution biases in the Lowe indices calculated with different frequency of basket-updates, is far more complicated than what has been discussed here, as it is also affected by the interaction of the different time periods involved in the calculation. Despite this, from this simple example, we can still find that the impact of the frequency of basket updates on the upper-level commodity-substitution bias depends on the relationship between a country's price trend and the expenditure pattern of different time periods. Intuitively, the more frequently the basket is updated, the more up-to-date baskets employed in the index calculation would be. In addition, if long-term price trends are persistent and consumers' price-induced commodity-substitution behaviour is as expected, increasing the frequency of basket updates would have an even larger possibility of lowering the commodity-substitution bias.

4.1.2 Empirical Results: Impact of the Basket-update Frequency on the Canadian CPI

Using the constructed data set, we compiled different CPI series by assuming different frequencies of updating the CPI basket while fixing the implementation lag equal to 13 months. Figure 4.1 shows the CPI series constructed with different frequencies of basket updates—from every year to every five years, and also with no basket update at all, for the period from January 2002 to December 2013.

Figure 4.1: Comparisons among the CPI series compiled with different frequencies of updating the CPI basket (January 2002=100[28])

[28] The January 2002 is the price reference period in the calculation.



Series "Freq_*x*" (*x*=1, 2,…5) in the above figure denotes CPI series compiled with frequency (*x*) of basket updating. It illustrates that the index level for a given time period gradually decreases as we increase the frequency of basket updates. The index level of the CPI series without basket updating is considerably higher than the other series. It is also noted that the index values do not obviously differ from each other within the first five or six years. For some months, the index values with basket-update frequency of every four years are even higher than those with basket-update frequency of every five years. We can observe more explicitly the impact of basket–updating frequency in Table 4.1 where we compare the annual chained-index level with the chain-linked Fisher index:

Table 4.1: Comparisons between different chained-CPIs, compiled with various frequencies of basket updating, and the chained-Fisher index (2003-2011)

	Indices (2003=100)	Difference in indices	Annual Growth Rate	Difference in Growth Rate
	2003 to 2011		(%)	(%)
Fisher	114.389	0.000	1.695	0.000
Lowe index-Every 1 year	115.857	1.468	1.857	0.162
Lowe index-Every 2 year	116.153	1.764	1.889	0.195

Lowe index-Every 3 year	116.547	2.157	1.932	0.238
Lowe index-Every 4 year	116.645	2.256	1.943	0.249
Lowe index-Every 5 year	116.918	2.528	1.973	0.278
Lowe index-No update	118.009	3.620	2.091	0.397

Examining these results, we find that the commodity-substitution bias could be reduced by increasing the frequency of updating the CPI basket during the examined period; however, the magnitude of the marginal reduction in commodity-substitution bias for each additional increase in the frequency of basket updating varied. If the frequency of updating the CPI basket was increased from every two years to every year in this exercise, the commodity-substitution bias, measured in terms of annual inflation rate, could be reduced from 0.195% to 0.162% on average; while, if the frequency of CPI basket updates was accelerated from every four years to every three years, the commodity-substitution bias would be reduced from 0.249% to 0.238%. Moreover, the impact was more significant when we changed the frequency from every four years to every two years; the substitution bias was reduced from 0.249% to 0.195% on average for the sample period. Thus, the numerical results in Table 4.1 lead to the conclusion that they are time sensitive, although in general, increasing the frequency of updating the CPI basket did reduce the commodity-substitution bias to certain degree.

4.2 Commodity-substitution Bias and the Implementation Lag of a New Basket

It is impossible to implement a new CPI basket in the basket reference period it refers to because of the time needed to conduct and process the Survey of Household Spending (SHS). This fact results in a certain time lag between the basket reference period and the implementation time of the basket. In this paper, this time lag is referred to as the implementation lag. It is widely recognized that shortening the lag of implementing a new CPI basket can lower the upward bias in a Lowe price index. In this section, we will revisit this common belief and verify how this lag influences the CPI.

4.2.1 Conceptual Impact of the Implementation Lag on the CPI

A specific example is used to show whether the commodity-substitution bias in the CPI could be reduced by shortening the implementation lag of a new basket.

If, for example, two baskets—the 2005 and the 2009 baskets—are available for the period from January 2009 to December 2012,[29] to implement the latter, a link month that chains indices across the two baskets is necessary.[30] We randomly select two possible link months for introducing the 2009 basket, specifically, December 2010 (which has a shorter implementation lag of 12 months) and April 2011 (which has a longer lag of 16 months). The difference in the CPI series calculated using these two possible link months[31] is then compared for assessing the impact of the implementation lag on the CPI.

For instance, the CPI from January 2009 to December 2012 using a shorter implementation lag—December 2010—as the link month, denoted by $P_{ChLo}^{2010,12}(2012,12)$, can be compiled as follows:

(32)
$$\begin{aligned} P_{ChLo}^{2010,12}(2012,12) &= P_{Lo}(p^{2009,01}, p^{2010,12}, q^{2005}) \times P_{Lo}(p^{2010,12}, p^{2012,12}, q^{2009}) \\ &= \frac{\sum_n p_n^{2010,12} \times q_n^{2005}}{\sum_n p_n^{2009,01} \times q_n^{2005}} \times \frac{\sum_i p_i^{2012,12} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}} \end{aligned}$$

The CPI for the same comparison periods using a longer lag—April 2011— as the link month, denoted by $P_{ChLo}^{L}(2012,12)$, can be compiled as follows:

(33)
$$\begin{aligned} P_{ChLo}^{2011,04}(2012,12) &= P_{Lo}(p^{2009,1}, p^{2011,04}, q^{2005}) \times P_{Lo}(p^{2011,04}, p^{2012,12}, q^{2009}) \\ &= \frac{\sum_n p_n^{2011,04} \times q_n^{2005}}{\sum_n p_n^{2009,01} \times q_n^{2005}} \times \frac{\sum_i p_i^{2012,12} \times q_i^{2009}}{\sum_i p_i^{2011,04} \times q_i^{2009}} \end{aligned}$$

The difference in the magnitude of the commodity-substitution bias in the two CPIs can be derived from the following expression:

(34)
$$\begin{aligned} &\left[P_{ChLo}^{2010,12}(2012,12) - P_{target}(2012,12)\right] - \left[P_{ChLo}^{2011,04}(2012,12) - P_{target}(2012,12)\right] \\ &= P_{ChLo}^{2010,12}(2012,12) - P_{ChLo}^{2011,04}(2012,12) \end{aligned}$$

---

[29] The introduction of 2009 basket is arbitrarily chosen as an example. January 2009, the price reference period, is the starting point of the 2009 basket reference period, the price observation period, December 2012, is arbitrarily chosen.
[30] In this simple example, a chain-linked Lowe index, defined in equation (11), will be calculated.
[31] Because of the inherent limitations of the Lowe formula, we believe that it will generate upward bias in most cases. Therefore, only upward bias will be taken into consideration.

Since the same target index is used to determine the difference in the magnitude of the commodity-substitution bias in the two CPI series, it is cancelled out in the comparison, and only the two CPIs calculated with different link months are left for conducting the comparison.

$$
\begin{aligned}
(35)\quad & P_{ChLo}^{2010,12}(2012,12) - P_{Ch-Lo}^{2011,04}(2012,12) \\
&= \left[\frac{\sum_i p_i^{2012,12} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}} \times \frac{\sum_n p_n^{2010,12} \times q_n^{2005}}{\sum_n p_n^{2009,01} \times q_n^{2005}} \right] - \left[\frac{\sum_i p_i^{2012,12} \times q_i^{2009}}{\sum_i p_i^{2011,04} \times q_i^{2009}} \times \frac{\sum_n p_n^{2011,04} \times q_n^{2005}}{\sum_n p_n^{2009,01} \times q_n^{2005}} \right] \\
&= \frac{\sum_i p_i^{2012,12} \times q_i^{2009}}{\sum_n p_n^{2009,01} \times q_n^{2005}} \times \left(\frac{\sum_n p_n^{2010,12} \times q_n^{2005}}{\sum_i p_i^{2010,12} \times q_i^{2009}} - \frac{\sum_n p_n^{2011,04} \times q_n^{2005}}{\sum_i p_i^{2011,4} \times q_i^{09}} \right) \\
&= \frac{\sum_i p_i^{2012,12} \times q_i^{2009}}{\sum_n p_n^{2009,01} \times q_n^{2005}} \times \frac{\sum_n p_n^{2010,12} \times q_n^{2005}}{\sum_i p_i^{2011,04} \times q_i^{2009}} \times \left(\frac{\sum_i p_i^{2011,04} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}} - \frac{\sum_n p_n^{2011,04} \times q_n^{2005}}{\sum_n p_n^{2010,12} \times q_n^{2005}} \right)
\end{aligned}
$$

A negative difference resulting from equation (35) would imply that a shorter implementation lag leads to a lower commodity-substitution bias. The last line of equation (35) indicates that the sign is determined by the difference of two prices indices, $\left(\frac{\sum_i p_i^{2011,04} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}} \right)$ and $\left(\frac{\sum_n p_n^{2011,04} \times q_n^{2005}}{\sum_n p_n^{2010,12} \times q_n^{2005}} \right)$. These latter two indices measure price changes between the two link months (December 2010 and April 2011) applying different baskets (the 2005 basket and 2009 basket). As mentioned before, generally speaking, price indices using a more obsolete basket tend to exceed those using a more up-to-date basket due to consumers' substitution behaviour. If this is the case, a negative sign should appear in the above difference, which leads to the conclusion that a shorter time lag would generate a lower bias, and the common belief would, therefore, be true. However, is this intuition always true? To verify this, the difference between these two indices is further examined.

To simplify the problem, we fix the items belonging to the two baskets.[32] Decomposing the index difference yields the following expression:

$$\frac{\sum_i p_i^{2011,04} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}} - \frac{\sum_i p_i^{2011,04} \times q_i^{2005}}{\sum_i p_i^{2010,12} \times q_i^{2005}}$$

(36) $$= \sum_i \frac{p_i^{2011,04}}{p_i^{2010,12}} \left(\frac{p_i^{2010,12} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}} - \frac{p_i^{2010,12} \times q_i^{2005}}{\sum_i p_i^{2010,12} \times q_i^{2005}} \right)$$

$$= \sum_i \frac{\overbrace{\left(\frac{p_i^{2011,04}}{p_i^{2010,12}} - P_{Lo}(p^{2010,12}, p^{2011,04}, q^{2009}) \right)}^{price\ variation} \times \overbrace{\left(\frac{q_i^{2009}}{q_i^{2005}} - Q_{Lo}(p^{2010,12}, q^{2005}, q^{2009}) \right)}^{quantity\ variation}}{Q_{Lo}(p^{2010,12}, q^{2005}, q^{2009})} \times s_i^{2010,12:2005}$$

where the Lowe quantity index is defined as:

(37) $$Q_{Lo}(p^{2010,12}, q^{2005}, q^{2009}) = \frac{\sum_i p_i^{2010,12} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2005}},$$

and the hybrid expenditure shares are defined as:

(38) $$s_i^{2010,12:2005} = \frac{p_i^{2010,12} \times q_i^{2005}}{\sum_i p_i^{2010,12} \times q_i^{2005}}$$

Thus, equation (36) demonstrates that the link month that yields lower commodity-substitution bias is determined by both price and quantity variations. It is, however, not easy to determine the sign of equation (36), because the price and quantity deviations are for two different periods. If the deviations in both prices and quantities are for the same period, it could be regarded as the

[32] In general, we would expect the items that belong to the different baskets to be different. This might give us slightly different price indices at the basic class level. Therefore, the comparison using official data is influenced by the difference in both the price vector and the quantity vector.

covariance between price relatives and the corresponding quantity relatives. In typical consumer theory, this covariance is negative—the price deviation $\left(\frac{p_i^{2009}}{p_i^{2005}} - P_{Lo}(p^{2005}, p^{2009}, q^{2009})\right)$ and the quantity deviation $\left(\frac{q_i^{2009}}{q_i^{2005}} - Q_{Lo}(p_i^{2010,12}, q^{2005}, q^{2009})\right)$ are negatively correlated. If the price trends for the period between the two possible link months (December 2010 and April 2011), represented by $\left(\frac{p_i^{2011,04}}{p_i^{2010,12}} - P_{Lo}(p^{2010,12}, p^{2011,04}, q^{2009})\right)$ is, on average, in the same direction as those between the two basket reference years (2005 and 2009), then we would expect that $\frac{\sum_i p_i^{2011,4} \times q_i^{2005}}{\sum_i p_i^{2010,12} \times q_i^{2005}}$ exceeds $\frac{\sum_i p_i^{2011,4} \times q_i^{2009}}{\sum_i p_i^{2010,12} \times q_i^{2009}}$. As a result, shortening the implementation lag could reduce the commodity-substitution bias.

In summary, this simplified case shows that a shorter implementation lag is associated with lower commodity-substitution bias as long as (i) the price trends between the two basket reference years are in the same direction as those between the two possible link months, and (ii) typical price-induced consumers' commodity-substitution behaviour is present. Price trends between the two basket reference years, in general, represent long-term price movements, whereas the price trends between two possible link months, if not too far from each other, normally reflect unpredictable price changes that are not necessarily in line with the long-term price movements, especially considering seasonal items. This implies that the impact on the CPI of shortening the implementation lag is not predictable. It depends on the consistency between the long-term price trends and short-term price fluctuations, and on consumer's commodity-substitution behaviour being as expected.

4.2.2 Empirical Results: Impact of the Implementation Lag on the Canadian CPI

In the first part of this section, we used the official CPI baskets without any adjustments and the associated CPI series to examine whether shortening the implementation lag for introducing the 2005 basket, the 2009 basket and the 2011 basket could reduce the commodity-substitution bias in the Canadian CPI.

The 2005 CPI basket was officially implemented in May 2007. To assess the impact of introducing the 2005 basket earlier than the official date on the commodity-substitution bias, we assume that it could have been implemented in any month from January 2007[33] to April 2007. A negative difference would be expected in the fifth column of Table 4.2 if introducing the 2005 basket earlier than May 2007 would have reduced the commodity-substitution bias. The results reported in Table 4.2 imply that implementing the 2005 basket earlier than May 2007 would not have yielded a lower CPI bias.

Table 4.2: Different link months for introducing the 2005 CPI basket

Possible Introduction month	**Possible Link Month**	$\frac{\sum_n p_n^{2007,4} \times q_n^{2005}}{\sum_n p_n^{link} \times q_n^{2005}}$ **(A)**	$\frac{\sum_i p_i^{2007,4} \times q_i^{2001}}{\sum_i p_i^{link} \times q_i^{2001}}$ **(B)**	**Difference (A)-(B)**
		2005 basket	**2001 basket**	
January 2007	December 2006	102.0116	101.9890	0.0226
February 2007	January 2007	101.9928	101.9447	0.0481
March 2007	February 2007	101.2472	101.2451	0.0021
April 2007	March 2007	100.3856	100.3813	0.0043

The sign of the difference between $\frac{\sum_n p_n^{2011,4} \times q_n^{2009}}{\sum_n p_n^{link} \times q_n^{2009}}$ and $\frac{\sum_n p_n^{2011,4} \times q_n^{2005}}{\sum_n p_n^{link} \times q_n^{2005}}$ listed in the fifth column of Table 4.3 determines whether the commodity-substitution bias would have decreased or increased by introducing the 2009 basket earlier than May 2011. Similarly, the sign of the difference between $\frac{\sum_n p_n^{2013,1} \times q_n^{2011}}{\sum_n p_n^{link} \times q_n^{2011}}$ and $\frac{\sum_n p_n^{2013,1} \times q_n^{2009}}{\sum_n p_n^{link} \times q_n^{2009}}$ in the fifth column of Table 4.4 determines whether the commodity-substitution bias would have decreased or increased by shortening the implementation lag of the 2011 basket. Conducting the same exercises for the 2009 CPI basket and the 2011 basket, we found that implementing the 2009 basket or the 2011

[33] Under current operational constraints, it takes at least 10 or 11 months to receive the SHS data to construct the CPI weights. Thus, the 2005 CPI basket could not be implemented earlier than January 2007.

basket earlier, under the constraint of the timeline of the SHS, would not have yielded a lower CPI bias. [34]

Table 4.3: Different link months for introducing the 2009 CPI basket

Possible Introduction month	Possible Link Month	$\frac{\sum_n p_n^{2011,4} \times q_n^{2009}}{\sum_n p_n^{link} \times q_n^{2009}}$ (A)	$\frac{\sum_i p_i^{2011,4} \times q_i^{2005}}{\sum_i p_i^{link} \times q_i^{2005}}$ (B)	Difference (A)-(B)
		2009 basket	**2005 basket**	
January 2011	December 2010	102.0339	102.0011	0.0329
February 2011	January 2011	101.7826	101.7540	0.0287
March 2011	February 2011	101.4966	101.4701	0.0266
April 2011	March 2011	100.4137	100.4076	0.0062

Table 4.4: Different link months for introducing the 2011 CPI basket

Possible Introduction month	Possible Link Month	$\frac{\sum_n p_n^{2013,1} \times q_n^{2011}}{\sum_n p_n^{link} \times q_n^{2011}}$ (A)	$\frac{\sum_i p_i^{2013,1} \times q_i^{2009}}{\sum_i p_i^{link} \times q_i^{2009}}$ (B)	Difference (A)-(B)
		2011 basket	**2009 basket**	
January 2013	December 2012	100.0678	100.0567	0.0111
February 2013	January 2013	100.0000	100.0000	0.0000
March 2013	February 2013	98.8240	98.8067	0.0173

In the following part of this section, the different CPI series calculated with different implementation lags using the constructed data set are reported. To isolate the impact of this phenomenon as opposed to the impact of basket–updating frequency, we fix the frequency of updating baskets at every 2 years, and vary only the implementation lag—between 12 months and 21 months. We also show the results of using one month as the implementation lag, which is operationally impossible under the current time line of the SHS, since the finalized expenditure data, taken mainly from the SHS, can be obtained only as early as ten months after the weight reference year.

[34] Similar exercises were performed outside of the time line of the SHS, occasionally, a negative sign was present. In addition, it could not be implied that the official link months were the optimal months for introducing the corresponding new CPI baskets because of the constraints on data availability.

Figure 4.2: Different CPI series corresponding to various implementations lags

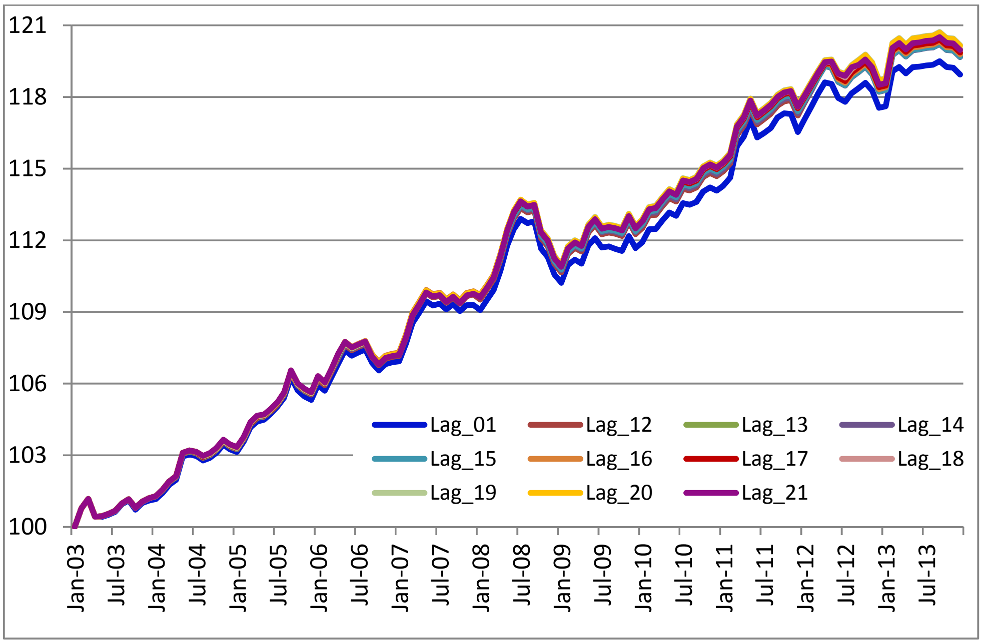


Figure 4.2 shows the cumulative impact of the implementation lags, which are kept unchanged for each CPI series, on the index values. In general, there are minor differences in the index values when the implementation lags are not significantly different from each other (for example, a 12-month lag versus a 13-month lag). This finding explains why many of the series cannot be distinguished separately in the above figure. However, over time, the series with longer implementation lags clearly begin to diverge from the series with shorter lags (for example, 18 months versus 12 months). It can also be demonstrated by the fact that the CPI series with a one-month implementation lag is significantly lower than the other CPI series.

Table 4.5 displays the comparison between the annual index levels of different CPI series corresponding to different implementation lags and the Fisher index series, which is highlighted in green. The lowest index value of each year, other than the Fisher index, is highlighted in red and the highest value is highlighted in yellow.

Table 4.5: Comparison of the different CPI series using various implementation lags with the Fisher index at the annual index level (2003=100)

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Fisher	100	101.728	103.746	105.475	107.401	109.624	109.670	111.404	114.389		

Lag_12	100	101.886	104.152	106.170	108.327	110.870	111.165	112.963	115.980	117.837	118.896
Lag_13	100	101.896	104.163	106.234	108.393	111.001	111.297	113.132	116.153	117.941	119.001
Lag_14	100	101.908	104.176	106.213	108.368	110.955	111.249	113.059	116.075	117.819	118.875
Lag_15	100	101.917	104.187	106.240	108.398	110.997	111.293	113.138	116.164	117.845	118.888
Lag_16	100	101.925	104.198	106.290	108.463	111.119	111.434	113.273	116.300	117.958	118.994
Lag_17	100	101.923	104.194	106.247	108.399	111.081	111.409	113.253	116.282	118.002	119.069
Lag_18	100	101.931	104.209	106.256	108.404	111.118	111.468	113.311	116.340	118.113	119.218
Lag_19	100	101.934	104.215	106.260	108.406	111.131	111.492	113.368	116.432	118.239	119.378
Lag_20	100	101.944	104.238	106.284	108.430	111.137	111.473	113.349	116.413	118.217	119.353
Lag_21	100	101.946	104.246	106.267	108.362	111.067	111.401	113.279	116.348	118.112	119.169

From 2003 to 2013, with the assumption of updating baskets every two years, we used the 2000, 2002, 2004, 2006, 2008 and 2010 baskets to calculate the CPIs. The implementation lag of introducing a basket was unchanged for each chained-CPI series during the examination period. Table 4.5 shows the long-term effect of introducing new baskets using a given implementation lag. For the years from 2004 to 2011, using 12 months as implementation lag resulted in the lowest annual index value; whereas for the year of 2012 and 2013, using 14 months as implementation lag yielded the lowest index value. In addition, using 21 months as implementation lag, the longest time-lags in the example, did not always yield the highest annual index values. We also found that the differences in index levels calculated using from12 months to 15 months as implementation lags were very small.

Table 4.6: Comparison of the index values and the associated geometric average growth rates of the different CPI series using various implementation lags and the Fisher index

	Indices (2003=100)	Differences in indices	Annual Growth Rate	Difference in Growth Rate
	2003-2011		(%)	(%)
Fisher	114.389		1.695	0.000
Lowe index, 1 month lag	115.484	0.000	1.816	0.121
Lowe index, 12 month lag	115.980	1.095	1.870	0.176
Lowe index, 13 month lag	116.153	1.591	1.889	0.195
Lowe index, 14 month lag	116.075	1.764	1.881	0.186
Lowe index, 15 month lag	116.164	1.686	1.891	0.196
Lowe index, 16 month lag	116.300	1.775	1.905	0.211
Lowe index, 17 month lag	116.282	1.911	1.903	0.209
Lowe index, 18 month lag	116.340	1.893	1.910	0.215

Lowe index, 19 month lag	116.432	1.951	1.920	0.225
Lowe index, 20 month lag	116.413	2.043	1.918	0.223
Lowe index, 21 month lag	116.348	2.023	1.911	0.216

Table 4.6 shows the comparison between the chained Fisher index and the chained price indices compiled using different implementation lags for the period from 2003 to 2011, as well as the corresponding geometric average growth rates. It indicates how the index values and the average inflation rate change with the implementation lags. Among the chained-CPI series that can be compiled in a timely manner, using 12 months as the implementation lag yielded the lowest inflation rate, however, the difference in the average inflation rate between using 12 months and 14 months as the implementation lag was only 0.01% for the sample period. As expected from the conceptual framework, the impact of the implementation lag on the CPI is not predictable, especially when we shorten or increase the lags by increments of 1 or 2 months. However, the commodity-substitution bias could still be reduced if the implementation lag were substantially shortened. This can be shown from the difference in the growth rates between using 1 month and 12 months as implementation lags.

From these empirical results, we cannot infer the impact on the CPI of a given link month of a particular CPI basket. To identify and illustrate this impact, the introduction of a specific CPI basket was examined. If, for example, the 2010 basket could be implemented between January 2012 and October 2012, any month from December 2011 to September 2012 could, therefore, be chosen as the link month. Using equation (36), we can determine retrospectively which month is the optimal link month for introducing the 2010 CPI basket. Table 4.7 shows the comparison between using April 2012 and all the other months, which are within the timeline of the SHS, as possible link months:

Table 4.7: Different link months for introducing the 2010 CPI basket

Possible introduction month	**Possible link month**	$\frac{\sum_n p_i^{2012,4} \times q_i^{2010}}{\sum_n p_i^{link} \times q_i^{2010}}$ **(A)**	$\frac{\sum_i p_i^{2012,4} \times q_i^{2008}}{\sum_i p_i^{link} \times q_i^{2008}}$ **(B)**	**Difference (A)-(B)**	**Order of magnitude of Difference**
		2010 basket	**2008 basket**		
January 2012	December 2011	101.779	101.582	0.197	8
February 2012	January 2012	101.322	101.188	0.134	6
March 2012	February 2012	100.864	100.770	0.094	4
April 2012	March 2012	100.399	100.370	0.029	2

May 2012	April 2012	100.000	100.000	0.000	1
June 2012	May 2012	100.054	99.975	0.079	3
July 2012	June 2012	100.556	100.402	0.154	7
August 2012	July 2012	100.687	100.477	0.210	10
September 2012	August 2012	100.380	100.175	0.205	9
October 2012	September 2012	100.200	100.094	0.107	5

We obtained positive differences in the fifth column of Table 4.7, implying that using months either earlier or later than April 2012[35] as the link month did not reduce the commodity-substitution bias in the CPI based on equation (36). To verify the results predicted by Table 4.7, we furthermore examined the monthly indices compiled using the above possible link months for introducing the 2010 basket:

Table 4.8: Comparisons of the monthly indices compiled using different implementation lags for the 2010 basket (January 2011-December 2012)

	Lag_12 (Dec-11)	Lag_13 (Jan-12)	Lag_14 (Feb-12)	Lag_15 (Mar-12)	Lag_16 (Apr-12)	Lag_17 (May-12)	Lag_18 (Jun-12)	Lag_19 (Jul-12)	Lag_20 (Aug-12)	Lag_21 (Sep-12)
Jan-11	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
Feb-11	100.281	100.281	100.281	100.281	100.281	100.281	100.281	100.281	100.281	100.281
Mar-11	101.313	101.313	101.313	101.313	101.313	101.313	101.313	101.313	101.313	101.313
…	…	…	…	…	…	…	…	…	…	…
May-12	103.787	103.723	103.683	103.616	103.586	103.668	103.668	103.668	103.668	103.668
Jun-12	103.269	103.206	103.166	103.099	103.069	103.151	103.228	103.228	103.228	103.228
Jul-12	103.135	103.071	103.032	102.964	102.935	103.017	103.093	103.150	103.150	103.150
Aug-12	103.450	103.387	103.347	103.279	103.250	103.332	103.409	103.466	103.462	103.462
Sep-12	103.636	103.573	103.532	103.465	103.435	103.517	103.595	103.652	103.648	103.546
Oct-12	103.830	103.767	103.726	103.659	103.629	103.711	103.789	103.846	103.842	103.740
Nov-12	103.548	103.485	103.445	103.377	103.348	103.429	103.507	103.564	103.560	103.458
Dec-12	102.912	102.849	102.809	102.742	102.712	102.794	102.870	102.927	102.923	102.822

In this table, the indices highlighted in green are identical because the same CPI basket (the 2008 basket) is used for the different CPI series. From September 2012, the index values are lower when the color of the cell gets darker. As implied by Table 4.7, the lowest index value should be

[35] If other months were chosen as the benchmark to compare with, both positive and negative signs would have shown up in the fifth column of Table 4.7.

presented in the CPI series with April 2012 as the link month. The numerical results reported in Table 4.8 verified what have been predicted by Table 4.7.

Although using April 2012 as the link month to introduce the 2010 basket generated the lowest index value, it might not necessarily be true for introducing other new baskets. Table 4.9 shows the comparison of index levels using different link months, which are in line with the time line of the SHS, to introduce the 2008 basket:

Table 4.9: Comparisons of the monthly indices compiled using different implementation lags for the 2008 basket (January 2009-December 2011)

	Lag_12 (Dec-09)	Lag_13 (Jan-10)	Lag_14 (Feb-10)	Lag_15 (Mar-10)	Lag_16 (Apr-10)	Lag_17 (May-10)	Lag_18 (Jun-10)	Lag_19 (Jul-10)	Lag_20 (Aug-10)	Lag_21 (Sep-10)
Jan 2009	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
Feb 2009	100.691	100.691	100.691	100.691	100.691	100.691	100.691	100.691	100.691	100.691
Mar 2009	100.921	100.921	100.921	100.921	100.921	100.921	100.921	100.921	100.921	100.921
…	…	…	…	…	…	…	…	…	…	…
May 2010	102.805	102.836	102.811	102.849	102.840	102.846	102.846	102.846	102.846	102.846
Jun 2010	102.679	102.711	102.686	102.724	102.714	102.721	102.718	102.718	102.718	102.718
Jul 2010	103.159	103.191	103.166	103.204	103.195	103.201	103.199	103.258	103.258	103.258
Aug 2010	103.097	103.129	103.104	103.142	103.132	103.139	103.136	103.196	103.196	103.196
Sep 2010	103.208	103.240	103.215	103.253	103.243	103.250	103.248	103.307	103.307	103.317
Oct 2010	103.604	103.636	103.611	103.649	103.639	103.646	103.643	103.703	103.703	103.713
Nov 2010	103.755	103.787	103.762	103.800	103.790	103.797	103.794	103.854	103.854	103.864
Dec 2010	103.637	103.669	103.644	103.683	103.673	103.680	103.677	103.737	103.737	103.746

Table 4.9 indicates that using December 2009 as the link month for introducing the 2008 basket yielded the lowest index levels. However, the differences in the index level using December 2009 or February 2010 as the link month are fairly small. Moreover, the results in both Table 4.8 and Table 4.9 lead to the conclusion that the optimal link month, which yields the lowest commodity-substitution bias, may differ from one new CPI basket to another.

In summary, the empirical results illustrate that the impact of shortening the implementation lag on the commodity-substitution bias is not predictable. Recently, Statistics Canada used 13 months as the implementation lag to introduce the 2011 basket because of operational constraints. The empirical results from this study suggest that shortening the lag to 12 months may have an insignificant impact on reducing the commodity-substitution bias. In addition, the link month that yields the lowest commodity-substitution bias may not always be the same caused by different monthly price fluctuations over time. As a result, it is not meaningful to fix the link month of implementing a new basket just for the purpose of reducing the commodity-substitution bias; besides, the optimal link month of introducing a new CPI basket cannot be determined in advance. However, since Statistics Canada also compiles the CPI annual table based on the calendar year, we recommend that the new baskets be introduced in January to have a consistent annual index.

5. Conclusion

The Lowe index number formula, one of the fixed-basket concept indices, is widely used by statistical agencies to compile their Consumer Price Indices (CPIs). However, due to its limitations associated with the fixed-basket concept, some concern arises from the use of this formula, in particular the issue of commodity-substitution bias. Because of the great importance of the CPI to its different users (such as central banks, policy makers, and the general population as a whole), researchers have devoted, and continue to devote, much work into investigating the issue of commodity-substitution bias in the CPI.

In this paper, we constructed a comprehensive data set by using two sources of information from Statistics Canada. Specifically, we used information from the Survey of Household Spending (SHS) for the years 2000 to 2011 and monthly CPI data for Canada at the basic class level for the period from January 2000 to December 2013.

This study focuses on the investigation of approaches to reducing the commodity-substitution bias in the Canadian CPI based on two key aspects associated with the introduction of new CPI baskets. Namely, (i) updating the CPI basket more frequently, and (ii) introducing new CPI baskets in a timelier manner. The empirical results indicate that increasing the frequency of updating the CPI basket can reduce the commodity-substitution bias. However, there is a diminishing marginal gain from further increasing the frequency of basket updates. The gains from changing the basket-update frequency from every four years to every two years are more significant than those from increasing the basket-update frequency from every two years to every year.

The impact of shortening the implementation lag for introducing a new CPI basket on the commodity-substitution bias is unpredictable because it depends on the consistency between the long-term price trends (between the two basket reference periods) and the short-term price movements (between the possible link months), as well as the existence of price-induced consumers' commodity-substitution behaviour. We found that clear differences can be observed in the price indices compiled by using 12-month implementation lag versus 21-month implementation lag, while the differences in the indices are small when comparisons are made with between 12, 13 and 14-month implementation lags. Therefore, based on both the decomposition of index differences and the empirical results, it is worthwhile for a statistical agency to pursue ways to dramatically shorten its implementation lag; however, small improvements to an already short implementation lag may not provide meaningful returns.

In this paper, we presented the empirical results using Canadian data. These results did not provide direct answers for choosing the most effective approach to reducing the commodity-substitution bias in a CPI. Statistical agencies in other countries can draw some inferences from these empirical works but should be cautious in generalizing these results to other CPIs because of the time and region dependence of the empirical results.

A. Appendix

The following table summarizes which basic classes required special methods of calculation, their corresponding CPI classification code, and the years which required these special calculations.

Table A.1: Basic classes required special methods of calculation

Basic Class	Classification Code	Year Affected
Basic classes below the aggregate level of Food and Non-Alcoholic Beverages Purchased in Stores	11	2000, 2002-2004, 2006-2008,
Basic classes below the aggregate level of Food and Non-Alcoholic Beverages Purchased from Restaurants	12	2000, 2002-2004, 2006-2008,
Mortgage Interest Cost	2201	2000, 2002-2004, 2006-2008, 2010
Replacement Cost	2202	2000, 2002-2004, 2006-20082010
Other household services	310505	2009- 2011
Financial services	310506	2009-2011
Women's Clothing	4101	2002
Men's Clothing	4102	2002
Children's Clothing	4103	2002
Women's Footwear (excluding athletic)	4201	2002-2004
Men's Footwear (excluding athletic)	4202	2002-2004
Children's Footwear (excluding athletic)	4203	2002-2004
Athletic Footwear	4204	2002-2004
Multipurpose Digital Devices	71010302	2009-2011

Generally, the backward and forward price updating procedure was implemented in order to construct weights for most of the above classes.

Several product categories were added to the 2009 basket. In order to align with the 2005 basket update classification, special adjustments such as excluding some sub-classes from its corresponding basic classes were made. For instance, from Other household services the following sub-class expenditures were excluded: legal services not related to the dwelling, funeral services, government services and retail club memberships. Similarly from Financial services, stocks and bond commissions were excluded.

The following table summarizes which price indices required special methods of calculation, their corresponding CPI classification code, and the years which required these special calculations.

Table A.2: Price indices required special treatments

Basic Class	Classification Code	Year Affected
Other household services	310505	Jan 2009 – Dec 2013
Financial services	310506	Jan 2009 – Dec 2013
Household Furnishings and Equipment, NES	3299	Jan 2000 – Apr 2007
Clothing	4101,4102,4103 4201,4202,4203 4204,430101,430102 4302,4303,4401,4402 4403,4404	Jan 2000 – Dec 2002
Other Public Transportation, NES	5299	Jan 2000 – Apr 2007
Eye Care Goods	610102	Jan 2000 – Apr 2007
Other Health Care Goods, NES	610103	Jan 2000 – Apr 2007
Eye Care Services	610201	Jan 2000 – Apr 2007
Audio Equipment	710301	Jan 2000 – Apr 2007
Cultural and Recreational Services, NES	710599	Jan 2000 – Apr 2007
Reading Material and Printed Material, NES	720299	Jan 2000 – Apr 2007
Other Alcoholic Beverages Purchased in Stores, NES	810299	Jan 2000 – Apr 2007

Special treatments are necessary for some classes in order to align with the 2005 basket update classification. For Clothing, higher level aggregated price indices were used as official baskets prior to 2005 contained significantly more detail. Similarly, for the *not elsewhere specified* (NES) classes, price movements were imputed from higher level indices.

Sub-class price indices were excluded from the following basic classes: Other household services and Financial services, as these did not exist in the 2005 basket classification structure. The same adjustment occurred with the expenditure weights.

Eye care goods and Eye care services are a combined class prior to the 2005 basket classification. For the years prior to 2005, the price indices for the combined class were therefore used for these two classes.

References

Diewert, W. Erwin. 1976. “Exact and superlative index numbers”, *Journal of Econometrics*. May. p. 115-145.

Généreux, Pierre A. 1983. “Impact of the choice of formulae on the Canadian Consumer Price Index.” *Price Level Measurement: Proceedings From a Conference Sponsored by Statistics Canada*. Erwin Diewert and Claude Montmarquette (eds.).

Greenlees, John S. and Elliot Williams. 2009. *Reconsideration of Weighting and Updating Procedure in the US CPI.* Paper presented at the 11th Meeting of the International Working Group on Price Indexes (Ottawa group) in Neuchâtel, Switzerland.

Ho, Ricky, Peter Campion and Chris Pike. 2011. *New Zealand Consumer Price Index—an empirical analysis of the frequency and level of weight updates.* Room document at the 12th Meeting of the International Working Group on Price Indexes (Ottawa group) in Wellington, New Zealand

International Labour Office. 2004. “Consumer Price Index Manual: Theory and Practice.” Geneva.

Rossiter, James. 2005. *Measurement bias in the Canadian Consumer Price Index.* Bank of Canada Working Paper 2005-39. December.

Statistics Canada. 2014 “The Canadian Consumer Price Index Reference Paper.” Catalogue No. 62-553-X.

Sabourin, Patrick. 2012. “Measurement bias in the Canadian Consumer Price Index: An update.” *Bank of Canada Review*. Summer.