Estimating Bias and Lag on Indexes Proposed To Be Moved To Less than Bimonthly Pricing

Onimissi M. Sheidu

U.S. Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Room 3650, Washington, DC 20212

Abstract

The CPI is a monthly measure of the average change in the prices paid by urban consumers for a market basket of consumer goods and services. It is constructed from approximately 8,000 basic indexes, which correspond to 38 geographic (87 PSUs) areas and 211 item categories. Although the prices of these items for the construction of the CPI are collected regularly, (usually on a monthly or bimonthly basis), not all these prices change at the same rate or even exhibit change. It has been suggested that some of the items with slower rate of change be moved to a less than bimonthly period of pricing. This study tries to assess and determine the existence of bias and lag for indexes with less than bimonthly pricing. The result of this study shows a very limited presence of small bias and lag on indexes of very few item strata that is not widespread either in scope or intensity for the pricing frequencies proposed and the index areas involved.

Key Words: Consumer Price Index, Price Relatives, Pricing Frequency, Bias, Lag.

Disclaimer: Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

1. Introduction

The CPI is a monthly index of all consumer items calculated for the total U.S. urban metropolitan and non-metropolitan population. The CPI is a weighted average constructed from approximately 8,000 basic indexes, which correspond to 38 geographic areas and 211 item categories, also known as item strata. An index area represents the most basic geographic area at which price indexes are calculated.

There has been a proposal to collect prices for certain items less frequently than they are currently collected. There are three index areas where all items are priced every month. In all other areas, most items have prices collected every other month. There are some exceptions, such as food at home items, which are collected every month. Analyzing the frequency of price changes has suggested that some items could be priced with reduced frequency. The objective of this study is to determine whether any bias or lag would be introduced by reducing the frequency with which prices are collected.

In order to study the issue of bias arising from reduced pricing frequency for all proposed changes in pricing cycle, it was necessary to start with monthly data. This limited the research to three self-representing Primary Sampling Units (PSUs) in which all quotes have prices collected monthly. These three PSUs are New York City (A109), Chicago (A207), and Los Angeles (A419). Having a monthly set of prices allowed simulations where prices were only collected every 2, 3, 4, or 6 months.

2. Brief Overview of CPI Calculation Method

In each of the 211 item strata there is one or more narrowly defined categories of goods and services called entry level items (ELIs). During the CPI item selection process, ELIs are selected from each stratum by a systematic probability proportional to size (pps) procedure, where the ELI weights are derived from expenditures reported in the two most recent years of the Consumer Expenditure Survey. ELI selections are independently drawn from each stratum for each sample replicate within each index area PSU. The selected ELIs are then priced in sample outlets on either monthly or bimonthly, or even on seasonal basis.

The CPI is constructed in two stages: in the first stage, which is known as the elementary cell stage, the price index for an item-area is computed every one or two months via a function of sample quote-level price changes called a price relative. For most commodities and services, the price relatives for the area-item combination are calculated using a weighted geometric mean. Thus the price relative for ELI j in item stratum i, is computed as:

$$X_{i,a}^t = X_{i,a}^{t-k} \times R_{i,a}^{t-k \to t} \text{, with}$$

$$R_{i,a}^{t-k \to t} = f(p_j^t, p_j^{t-k}; w_{jia}^b) = \prod_{j=1}^{S_{ia}} \left(p_j^t / p_j^{t-k} \right)^{w_{jia}^b}$$

where,

$X_{ia,t}$ is the current period index at time t, in item stratum i of index area a, relative to base period,

$X_{ia,t-k}$ is the index from the previous period, at time $t-k$, in item stratum i of area PSU a,

$R_{i,a}^{t-k \to t}$ is the ELI j price relative in item stratum i, for index area a, at the current time t, in relation to period $t-k$,

$P_{iaj,t}$ is the price for ELI j in item stratum i, for index area a, at the current period t,

$P_{iaj,t-k}$ is ELI price at time t-k in item stratum i, for index area a, S_{ia} is the sample for item stratum i containing item j in area a, and w_{jia}^b represents the quote-level sampling weight of sample ELI j, normalized to the same sample rotation base period reference month b for all quotes in the item-index area a.

In this study, one of the main objectives is to determine if there is a shift in short term price changes at the elementary cell level in form of bias or lag as a result of moving the designated item strata to a reduced frequency of price collection.

3. Study Design

The study is conducted at the Elementary Level Items (ELI) level within three self-representing Primary Sampling Units:
New York City – A109;
Chicago – A207;
Los Angeles – A419.

All items are priced monthly in these PSUs. By having a price every month, it is easier to simulate indexes based on a variety of pricing frequencies. The simulation study is conducted using monthly data from 89 ELIs for the period of January, 2004 through

December, 2008. Table 1.1 shows some of the ELIs which were proposed to move to a decreased frequency of price collection.

Although the main focus of this study is on indexes of item strata to be moved to less than bimonthly pricing, we have included item strata proposed to be moved from monthly to bimonthly pricing as a matter of supplementing the study to gain more insight into the overall impact of reduced pricing cycles.

Table1.1:   Examples of ELIs moving to a reduced pricing frequency

Examples of ELIs for Monthly to bimonthly Cycle		
Elementary Level Item (ELI)		Item Stratum
ELI	Description	
ED031	Cellular Telephone Service	ED03
FT051	Baby Food	FT05
RG011	Single-Copy Newspaper and Magazines	RG01
Examples of ELIs for Bimonthly to quarterly cycle		
FW021	Distilled Spirits at Home	FW02
MA011	Prescription Drugs	MA01
RB011	Pet Food	RB01
Examples of ELIs for Bimonthly to 4-monthly Cycle		
FV011	Full Service Meals And Snacks	FV01
GC011	Haircuts And Other Personal Care Services	GC01
HM011	Paint, Wallpaper Tools And Supplies	HM01
Examples of ELIs for bimonthly to 6-monthly Cycle		
GD051	Checking Account And Other Bank Services	GD05
MC021	Dental Services	MC02

Index simulations were run for each of the A-PSUs (A109, A207 and A419) and all of the 89 ELIs involved. The simulations were run for each month for a period of five years (January 2004 – December 2008) using both the normal ELI frequency of price collection and the proposed ELI frequency of price collection. The monthly price quotes were adjusted while leaving the price quotes of the remaining ELIs within the item stratum untouched. For instance, if an ELI is currently priced bimonthly and is proposed to move to quarterly pricing, then one simulation would be run as if the ELI was bimonthly (in a monthly PSU bimonthly ELIs are priced monthly) and another simulation would be run as if it were priced quarterly. The process entails substituting the prices of the ELIs of interest within every item stratum of PSUs: A109, A207, and A419 in a given month, t, with their corresponding previous month's price quotes based on the pricing cycle.

Take the example of HK022 (Small Electric Kitchen Appliances), which is priced bimonthly and has been proposed to be switched to quarterly pricing. In order to simulate an index based on quarterly pricing in July 2008 (assuming the quote is priced in that month) the price quotes for the previous months, t-1 and t-2, are replaced by the April 2008 price quote (the previous time the quote was priced). Price quotes in the remaining ELIs (HK021, Floor Cleaning Equipment, and HK023, Other Electrical Appliances) within the item stratum HK02 are not adjusted for calculation of the July 2008 index for the item stratum HK02. To produce a July 2008 index based on bimonthly pricing, the June (t-1) price quote is replaced with the May 2008 (t-2) for all HK022 quotes and other quotes are left unchanged.

Initially, indexes are simulated based on a bimonthly pricing frequency for all of the ELIs normally priced bimonthly for the period of January 2004 through December 2008. This provides the 'True Bimonthly data', based on the current CPI methodology. This is used as a benchmark to which all other simulation results are compared. Then simulations are run with prices adjusted to reflect the proposed pricing frequency. The simulations are only run for item strata where at least one ELI has been proposed to be changed to a reduced frequency of price collection as these are the only indexes where bias or lag could be introduced by reducing the frequency of price collection.

4. Calculated Statistics

4.1 Item Relatives and Index Percent Change
We calculate index percent changes for the item strata to enable the calculation of the index bias and lag.

Index percent change measures the degree of price change between any two periods.
To calculate index percent changes, we first had to calculate two types of price relatives per item stratum via weighted geometric means. Thus,

1). Unadjusted price relative - use normal bimonthly quotes, j:

$$R_{i,a}^{t-1 \to t} = f(p_j^t, p_j^{t-1}; w_{jia}^b) = \prod_{j=1}^{S_{ia}} \left(p_j^t \Big/ p_j^{t-1} \right)^{w_{jia}^b}$$

2). Adjusted price relative - use quotes j, adjusted to reflect proposed pricing cycles:

$$R_{i,a}^{t-k^* \to t} = f(p_j^t, p_j^{t-k^*}; w_{jia}^b) = \prod_{j=1}^{S_{ia}} \left(p_j^t \Big/ p_j^{t-k^*} \right)^{w_{jia}^b}$$

Here,
t-1 is the previous month, and *t-k** is the adjusted month *t-k,* whose ELI quotes is used,
p_j^t is the price for ELI *j* at the current month t,

$p_j^{t-k^*}$ is the price for ELI *j* at the adjusted month t-k*,

S_{ia} is the sample for item stratum *i* containing ELI *j* in area *a*

w^b_{jia} is the quote-level sampling weight of sample ELI j, normalized to the same sample rotation base period reference month b for all quotes in the item-index area a.

Two corresponding Price Indexes are computed by making use of the above price relatives:
Unadjusted Index:

$$X^t_{i,a} = X^{t-1}_{i,a} \times R^{t-1 \rightarrow t}_{i,a} \text{, and}$$

Adjusted index:

$$X^{t*}_{i,a} = X^{t-k*}_{i,a} \text{ x } R^{t-k* \rightarrow t}_{i,a}$$

With the above two different price indexes, which are calculated for every item strata involved in the study, we calculated the following two corresponding index Percent changes (PCs):
The Unadjusted Index Percent change (PC), which is calculated as:

$$PC^t_{i,a} = \left[\left(X^t_{i,a} - X^{t-1}_{i,a} \right) \Big/ X^{t-1}_{i,a} \right] x100$$

And the Adjusted Index Percent Change (PC*) computed as:

$$PC^{t*}_{i,a} = \left[\left(X^{t*}_{i,a} - X^{t-k*}_{i,a} \right) \Big/ X^{t-k*}_{i,a} \right] x100$$

4.2 Bias and Relative Bias
The results from each batch of simulations are used to calculate bias, relative bias and lag.

The bias is calculated as: the difference between the index percent changes calculated using quotes adjusted to reflect the proposed pricing cycles (Adjusted) and the calculated index percent changes based on the normal (Unadjusted) bimonthly quotes, as currently used in the majority of CPI index PSUs.

The *relative bias* expresses the bias as a proportion of the 'Unadjusted" index percent changes, that is, the index percent changes computed using the item strata index derived from bimonthly pricing quotes.

Thus,
Bias = Average difference between the adjusted and the unadjusted index percent changes (PC) for all the item strata within index PSU over the study period.
Bias and relative bias are calculated based on three periods (monthly, yearly, and five-year):

Monthly bias is calculated as: $Bias_{iAt} = PC^*_{(i,A,t)} - PC_{(i,A,t)}$

Monthly relative bias (RB): $RB_{iAt} = Bias_{iAt} \Big/ PC_{(i,A,t)}$, $a \in A$, A= all 3 PSUs.

The yearly bias and relative bias are calculated for each year as the yearly average of the monthly bias and relative bias for item strata index percent changes per index PSU. That is, yearly item strata index bias and relative bias are computed as:

$$Bias(year \delta)_{i, PSU_A} = Bias_{iAy} = \sum_{t=t0}^{t0+11} Bias_{iAt} / 12 , \text{ and}$$

$$Relative.Bias(year \delta)_{i, PSU_A} = RB_{iAy} = \sum_{t=t0}^{t0+11} RB_{iAt} / 12 .$$

Hence, the bias for the study period is:

$$Bias(2004\text{-}2008)_{i, PSU_A} = Bias_{iA} = \sum_{y=1}^{5} Bias_{iAy} / 5$$

While the relative bias is:

$$Relative.Bias(2004-2008)_{i, PSU_A} = RB_{iA} = \sum_{y=1}^{5} RB_{iAy} / 5$$

Where,

$year \delta$, y = Year: 2004*, 2005, 2006, 2007, 2008,

$PC^*(i, A, t)$ =item stratum index percent change derived from the proposed pricing cycle,

$PC(i, A, t)$ = item stratum index percent change derived from current CPI pricing cycle,

i = item stratum with proposed shift in pricing frequency,

a = PSU (A109 or A207 or A419),

A = All 3 PSUs (A109, A207, and A419) combined,

t = Current month.

2004* - monthly calculation only involves 10 months index percent change. Because of quotes adjustments, the calculation of the initial price relative starts from March, 2004.

The computation of relative bias gives a better assessment of the level of bias occurrence in every item stratum in each PSU, as well as among the PSUs. Every computation sequence noted above is done separately for each index PSU (A109, A207, and A419), and for all the 3PSUs combined.

4.3 Lag

The Lag is calculated as the difference between two series of smoothed monthly adjusted index percent changes and unadjusted index percent changes – the average over time of the difference between monthly index percent change derived from *adjusted* pricing cycle and that derived from *CPI* bimonthly pricing cycle.

The monthly item strata index percent changes from both adjusted and CPI bimonthly (unadjusted) pricing cycles are expressed as functions of time *t*.

In which, $\overline{y}^*_{i,a,t} = f_i^*(t)$ for the monthly adjusted index percent change for item

stratum *i*. And $\overline{y}_{i,a,t} = f_i(t)$ for the monthly unadjusted index percent change for item

stratum *i*. The average of their differences over the study period gives us the lag value for that period for item stratum *i*. In this case, it is assumed that smoothing the averages will neutralize the source of randomness.

Monthly lag for item stratum i: $lag_{i.a,t} = \bar{y}*_{i,a,t} - \bar{y}_{i,a,t}$, $\quad\quad i \in a$, $a \in A$

Monthly lag per PSU a, (all 75 item strata combined) is

$$lag_{a,t} = \sum_{i=1}^{75} lag_{i,a,t} / 75$$

Yearly lag per PSU (all 75 item strata combined): $lag_{a,y} = \sum_{t=1}^{12} lag_{a,t} / 12$

The 5-year lag for item stratum i, is calculated:
a). For every PSU a, as $lag_{i,\,2004-2008\,PSUi} = lag_{i,a} = \sum_{t=1}^{58} lag_{i.a,t} / 58$

b). for all 3 PSUs, A, combined as $lag_{i,A} = \sum_{a=1}^{3} lag_{i,a} / 3$

All estimates are calculated per pricing cycle.

5.   Analysis Result

5.1 Initial Analysis Result

The result of the analysis as shown in Table 1 below, revealed the presence of both upward (positive) and downward (negative) bias in the index percent changes of some of the item strata whose ELIs are proposed for a less frequent pricing. In the majority of cases, the bias is negative, which shows that the index percent change for an item stratum calculated as a result of moving it to a reduced pricing frequency is smaller than its index percent change calculated using the CPI bimonthly pricing cycle.

 In this study, we focused more on the overall result for the study period with all the 3 PSUs combined, rather than the analysis conducted on individual PSUs. The individual item strata per PSU analyses are conducted, but they serve as a means of zooming into the details of the changes, and thus help to better understand the broader overall result for the whole study period and for all the 3 PSUs combined. Besides, we examined month to month detail analysis of item strata with ELIs responsible for extreme bias and lag to get an assessment of the occurrences of the bias, whether the bias is serial or a single occurrence (aberration).

Table 1 shows average yearly bias and lag of combined item strata by pricing cycles for all the 3 PSUs combined. The table demonstrates more or less evenly distributed lag from year to year. In the table, there were very few cases of spike in lag, and they are mainly to the downside: -0.056 and -0.063 in 2005 and 2006 respectively in the Bimonthly to Quarterly pricing ELI group. This shows that the percent changes of item strata index calculated under current pricing frequency are larger than the same item strata index calculated using the proposed reduced pricing frequency. This result supports the more detail analysis of bias and lag done on individual item strata for each of the 3 PSUs (A109, A207, and A419).

In addition, table 1 also shows that the relative biases, as well as the lag in the item group proposed to move from bimonthly to quarterly pricing are higher, compared to in all other groups.

Also, while the values of relative bias and lag on the index increase in the later part of the study period in both the item strata group containing ELIs proposed for triannual pricing and the item strata group containing ELIs proposed for semi-annual pricing, they are more or less decreasing in both item strata groups containing ELIs proposed to move from monthly to bimonthly and bimonthly to quarterly pricing from year over year in most cases.

Table 1: Comparative analysis of yearly bias and lag for all the 75 item strata (combined) proposed to move to less pricing frequency for all three PSUs (A109, A207, A419 combined) between January 2004 and December 2008.

Year	Monthly to Bimonthly			Bimonthly to Quarterly			Bimonthly to Triannual			Bimonthly to Semi-annual		
	Absolute Bias	Relative Bias	Lag	Absolute Bias	Relative Bias	Lag	Absolute Bias	Relative Bias	Lag	Absolute Bias	Relative Bias	Lag
2004	-0.014	-0.048	-0.023	-0.025	-0.289	-0.037	-0.008	-0.083	-0.014	-0.001	-0.006	-0.001
2005	0.016	0.041	0.027	-0.030	-0.116	-0.056	0.006	0.025	0.007	-0.003	-0.015	-0.002
2006	0.001	0.003	0.001	-0.038	-0.244	-0.063	0.001	0.004	0.003	0.020	0.074	0.022
2007	-0.002	-0.003	-0.002	0.010	0.144	0.011	-0.009	-0.043	-0.016	-0.008	-0.039	-0.012
2008	0.001	0.001	0.001	0.015	0.077	0.028	-0.021	-0.058	-0.023	0.016	0.062	0.022

In table 2 that follows, all item strata under study are grouped into acceptable and unacceptable group based on the level of relative bias of their index percent changes. The acceptable group consists of 49 item strata with 57 ELIs, which is about 65% of all the item strata and ELIs listed for the study. The index percent changes for the item strata under this category also has zero to 0.05 relative biases. The analysis of lag also shows that most item strata in this group also have very small or no lag between the adjusted and unadjusted index percent changes. The indexes of the item strata in this group exhibit either a downward bias of between -0.05 and 0.0, or an upward bias (very few) of less than 0.05.

About 11 item strata have moderate level of relative biased of index percent change with downward relative bias of between -0.1 and -0.05, and upward relative bias of between 0.05 and 0.1. The remaining 15 item strata exhibit either extreme downward bias of below -0.1, or upward bias of above 0.1.

Items with moderate level of relative bias together with the ones with extreme level of relative bias formed the unacceptable item strata group, which is 35 % of all the items examined .This group are not suitable for the proposed reduced pricing frequency unless further evaluation is done. The lag between the adjusted and the unadjusted average index percent changes of these item strata is relatively higher compared to that of the item strata under the group with acceptable level of relative bias.

Table 2: Summary of item strata bias by level of relative bias by pricing cycle (2004-2008).

| *LEVEL OF BIAS*	Rel. Bias (absolute)	ITEM STRATA PRICING CYCLES								*ALL TOTAL*	
		Monthly to Bimonthly		Bimonthly to Quarterly		Bimonthly to 4-Monthly		Bimonthly to Semi-annual			
		Items	*In %*	*Items*	*In %*	*Items*	*In %*	*Items*	*In %*	*Total*	*In %*
LOW (ACCEPTABLE)	0.00 - 0.05	15 (16)	83 (84)	9 (12)	36 (40)	21 (25)	78 (74)	4 (4)	80	49 (57)	65
MODERATE (UNACCEPTABLE)	0.05 - 0.10	1 (1)	6 (5)	5 (5)	20 (17)	4 (7)	15 (20)	1 (1)	20	26 (31)	35
EXTREME (UNACCEPTABLE)	0.10 - Above	2 (2)	11 (11)	11 (13)	44 (43)	2 (2)	7 (6)	0 (0)	0		
TOTAL		18 (19)	100	25 (30)	100	27 (34)	100	5 (5)	100	75 (88)	100

The result of simulation analysis on monthly lag of the index percent change by proposed group of pricing frequency by PSU shows that there is slight lag in all the three PSUs. It is about -0.002 in Chicago, for item strata proposed to move from bimonthly to quarterly pricing. In New York the lag for this same group of item strata is -0.029 with relatively stable trend, while in Los Angeles it is about -0.01 with increasing divergence between "adjusted" percent change and the "unadjusted" percent change (see fig 1 through 3).
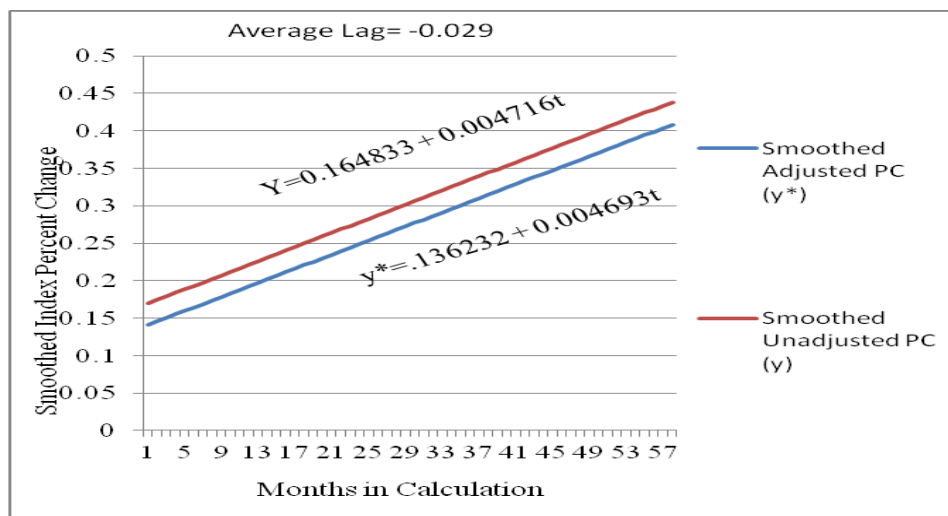
Figure 1: Smoothed monthly Percent Change for adjusted and unadjusted index for item strata to be moved from bimonthly to quarterly pricing in New York, 2004-2008.
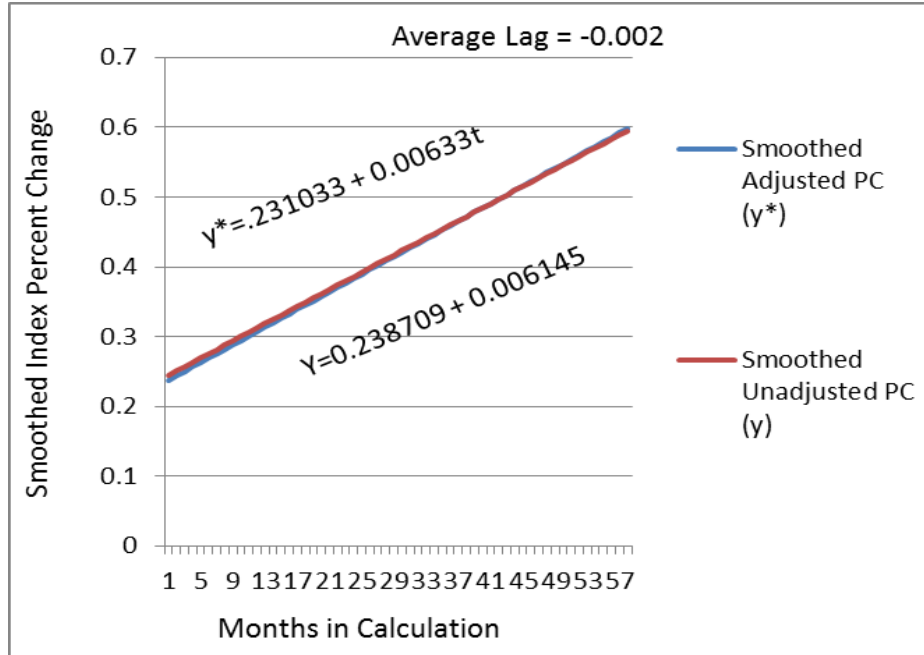
Figure 2: Smoothed monthly Percent Change for adjusted and unadjusted index for item strata to be moved from bimonthly to quarterly pricing in Chicago, 2004-2008.
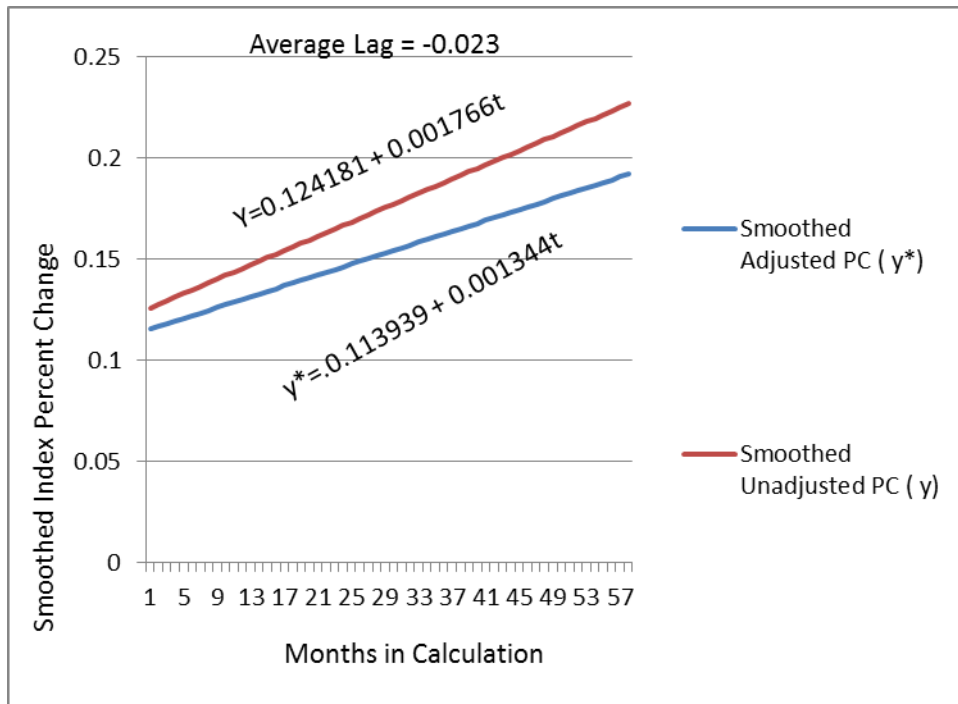
Figure 3: Smoothed monthly Percent Change for adjusted and unadjusted index for item strata to be moved from bimonthly to quarterly pricing in Los Angeles, 2004-2008.

It is important to note that the "adjusted" index percent change trails the "unadjusted" index percent change in most cases, especially among the item strata proposed for quarterly and triannual pricing. The lag for all PSUs combined for the study period merely reflects the result from an individual PSU in most cases.

5.2 Re-evaluation of 26 Item Strata with Extreme Bias and Lag
There are 26 Item strata with extreme bias and lag, which were reassigned to all 3 pricing groups (bimonthly to quarterly, bimonthly to triannual, and bimonthly to semi-annually).

Result shows improvement in the level of relative bias, as well as the lag of the item strata index percent changes. Over 50 % of the item strata have less than 0.05 (acceptable level) of relative bias. And all the 26 item strata have little (less than 0.05) or no lag.

6. Comments and Recommendations

Although this study shows some presence of bias and lag in the index percent change of some of the item strata proposed to be priced with reduced frequency, majority of the item strata, however have little or no relative bias and lag on their index percent change as a result of the switch in pricing frequency.

The result of the study makes the following findings possible:

1. Item strata can simply be mapped into two groups - the ones that are eligible for a possible move to a reduced pricing cycle and those that are not suitable (requiring further investigative studies)**.**

2. On average, item strata with little or no bias also tend to have smaller lag between adjusted index percent change and the unadjusted index percent change for the study period.

3. Although item strata index bias and lag for the three cities (Chicago, Los Angeles and New York) investigated are not homogeneous for the individual items across PSUs, they mostly trend in the same direction for the period investigated.

4. Many item strata proposed to move from bimonthly to quarterly pricing have high relative biases in contrast to item strata in other pricing groups, notably, bimonthly to triannual, and bimonthly to semi-annual pricing groups.

5. Most item strata have downward (negative) bias, meaning that the average index percent change calculated from bimonthly pricing is higher than the resulting average index percent change computed from the propose reduced pricing cycle.

6. When moving some item strata to a reduced pricing cycle, high bias and lag could be avoided by initially assigning all the item strata to all the reduced pricing groups (bimonthly to quarterly, bimonthly to triannual, and bimonthly to semi-annual), assessing their performance, and then choosing the pricing group with the smallest relative bias and lag for each specific item stratum.

In conclusion, the result of this study seems to support the idea to move some item strata with slower rate of price change to a reduced pricing frequency. But some of these items may require special investigation before any change to their pricing frequency could be implemented to curb unacceptable bias and lag on their indexes.

Further research could be conducted using the two PSUs comprising the New York suburbs (CPI codes A110 and A111) and the Los Angeles suburbs (CPI code A420)
 The research could also be extended in time.

Acknowledgements

References

Bureau of Labor Statistics (2003), BLS Handbook of methods, Washington, D.C
http://www.bls.gov/opub/hom/home.htm

Kish, Leslie (1965) Survey Sampling, Wiley, New York.

Lane, Walter (1996) "Changing the CPI Item Structure," U.S. Bureau of Labor Statistics
http://www.stats.bls.gov/mlr/cpiwl001.htm

Leaver, S. and Valliant, R. (1995) "Chapter 28: Statistical Problems in Estimating the U.S. Consumer Price Index," Business Survey Methods, Brenda G. Cox, et al., editors, Wiley, New York.

Ron Cody (2001) Longitudinal Data and SAS: A Programmer's Guide.

Groves, Robert (1990) Cost and Error Modeling in Social Science Surveys, Wiley, New York.

Hansen, Morris G., Hurwitz, William N., and Madow, William G. (1953) Sample Survey Methods and Theory, Wiley, New York.