Estimating Sampling Weights for the U.S. Consumer Price Index

Sylvia G. Leaver and Darin T. Solk
U.S. Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Room 3655, Washington, DC 20212-0002
Leaver_S@bls.gov

Keywords: Probability sampling, consumer expenditure, variance estimation

In this paper, we present findings from current research on quote-level weights used in constructing elementary indexes for the Commodities and Services (C&S) component of the U.S. Consumer Price Index. These weights are complex in their construction and rely on expenditure estimates from two supporting surveys, the Consumer Expenditure Survey and the Telephone Point of Purchase Survey, both of which contribute to the variability of the quote-level weights. We also present experimental results, including estimates of indexes and standard errors, for three major groups with which we use smoothed estimates of daily expenditures.

In Section One the official CPI price index estimator, sample design, and basic quote weight are described. Section Two presents a discussion of basic weight estimation, extreme quote weight analyses, and stratified jackknife variance estimation. Section Three presents computational results and compares estimates of the index computed with smoothed full-sample expenditure estimates. Sources of price change variability are identified and discussed. Conclusions are given in Section Four.

1. Background

The CPI is calculated monthly for the total U.S. metropolitan and urban non-metropolitan population for all consumer items, and it is also estimated at other levels defined by geographic area and item groups such as cereal, women's suits, and tobacco products (BLS, 2003).

An index area is the most basic geographic area for which a price index is computed. There are two types of index areas: self-representing areas, such as New York, which were selected with certainty; and non-self-representing areas, whose sample comprises two or more primary sampling units (PSUs) selected according to a probability sample. The 1998 revised U.S. All Cities CPI is a weighted average of 38 index area CPI's; 31 from self-representing and 7 from non-self-representing areas. For purposes of variance estimation and operational manageability, the sample for each self-representing PSU is segmented into two or more subsets, called sample replicates. The C&S sample is refreshed on a rotating basis with approximately one-quarter of the item and outlet sample in each PSU being reselected each year.

Each item stratum is composed of one or more narrowly defined classes called entry level items (ELIs). An ELI describes the level of specification for a class of goods with which a data collector enters an outlet for initial pricing. In CPI sample selection, 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 for each sample replicate within each PSU.

Sample frames and weights used in outlet selection are derived from the Telephone Point of Purchase Survey (TPOPS), a random-digit telephone survey conducted by the U.S. Bureau of the Census for BLS. TPOPS provides the names and addresses of outlets, and dollar amounts, of purchases for classes of items known as POPS categories. A POPS category is a group of items normally sold in the same kind of outlet. Each ELI belongs to only one POPS category. Outlet frames, total daily expenditure estimates, and selection probabilities are derived from POPS data for each PSU-POPS category-sample replicate.

In outlet selection, outlets are selected via systematic pps from frames for each PSU-sample replicate for POPS categories corresponding to ELIs selected in item sampling. Selected items are then priced in sample outlets on a monthly, bimonthly, or seasonal basis.

The CPI is estimated for items grouped into 211 strata for each index area, although not all such indexes are published every month. It is constructed in two stages. In the first or elementary cell stage, the price index for an item-area is updated every one or two months via a function of sample quote-level price changes called a price relative. Let X_{ia}^{t} denote the index at time t, in item stratum i, area a, relative to time period 0. Then

$X_{ia}^{t} = R_{ia}^{t,t-1} X_{ia}^{t-1}$ where $R_{ia}^{t,t-1}$ denotes the price relative between times t and t-1 . Since 1999, price relatives for most commodities and services, including all food and apparel commodities, have been computed using a weighted geometric average (BLS, 1997):

$$R_{ia}^{t,t-1} = \prod_{j \in S_{ia}} \left(P_{iaj,t} / P_{iaj,t-1} \right)^{w'_{iaj}} =$$

$$e^{\sum_{j \in S_{ia}} w'_{iaj} \ln \left(P_{iaj,t} / P_{iaj,t-1} \right)};$$

Here S_{ia} represents the sample for item i in area a, P represents the price, and w' represents the quote-level sampling weight of sample item j, normalized to the same sample rotation base for all quotes in the item-index area. This paper will focus on the quote-level sampling weight, which has the following structure:

$w_{iaj} = \alpha Efgr / MBn$, where

α is the percent of sales of the ELI to the total sales of the POPS category in the sample outlet;
E is the basic weight, an estimate of total daily expenditures for the POPS category, for the CPI urban population in the PSU and sample replicate. For geometrically averaged price relatives, this daily expenditure is normalized to the same reference period, December 1997, because the entire sample for any item stratum is not rotated at the same time for the three largest self-representing PSUs and for non-self-representing index areas;
f is a duplication factor which accounts for any special sub-sampling of outlets or quotes;
g is a geographic factor which accounts for differences in coverage of the PSU between index revisions;
M is the number of usable quotes for the stratum-PSU replicate;
r/n is the share of n stratum selections that were selected for the ELI; and
B is the probability of selection of the ELI within the stratum in the PSU-sample replicate, which is, in most cases, the proportion of the expenditure for the ELI of the total expenditure for its item stratum in its Census region.

2. Estimates of Quote Weight Terms

Estimates for each of the terms in the weight formula are derived from the sample surveys which support the index, the two principle sources being the Consumer Expenditure Survey and TPOPS. As currently computed, quote-level weights can vary remarkably among the observations within a basic price-index cell. Values of ratios of the most heavily weighted to least heavily weighted quotes within a basic cell range from less than 10 to greater than 100. An evaluation of active quotes in the C&S sample for the CPI for June 2002 through September 2003 revealed that the dominant terms affecting extreme quote weight ratios were the basic weight (E) and percent of POPS (α) terms. See Table 1 below.

Recently, an inter-divisional team in the CPI program at BLS has been investigating these estimates and is engaged in assessing their data sources with respect to appropriateness and reliability for use in the construction of the price relative. The focus of the research presented herein will be the term E, the estimate of total daily expenditures for a POPS category, currently estimated at the PSU and sample replicate level.

This estimate is derived from expenditures reported in the TPOPS household survey. Separate estimates are computed for each PSU and sample replicate. Variation in these estimates among PSUs and sample replicates and among PSUs in non-self representing index areas is the largest source of variation in elementary-cell quote weights in 40 percent of the elementary cells. Among the three major groups in this study, Food and Beverages, Apparel, and Housing, extreme variation among basic weights is the dominant contributor to extreme quote weight ratios in 47 percent of elementary cells.

For this study, an experimental data set was created from extractions from the Commodities and Services database for June 2002 to December 2003. The initial focus of the study was on the 108 priced item strata comprising the Food and Beverages, Apparel, and Housing C&S major groups. Most item strata in Food and Beverages are priced monthly in all PSUs. Apparel and Housing strata are more generally priced bimonthly in all but the three largest index areas. These three major groups contain 29 multi-ELI strata and 79 single-ELI strata. Price relatives for all but four housing strata are estimated using a geometric average (BLS,1997).

Price relative estimates were computed for July 2002 to December 2003 using production values of collected prices and quote-level weights, and then again using smoothed estimates for the normalized basic weights for each ELI.

Table 1. First and Second Largest Factors Associated with Extreme Final Quote Weight Ratios, U.S. CPI Commodities and Services Sample, June 2002-September 2003

	Largest Ratio														Either First or Second Largest Ratio	
	Basic Weight		Duplication Factor		ELI Factor		Geographic Factor		M Value Factor		% of POPS Factor		Total			
Second Largest Ratio	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Basic Weight	0	0.00	97	0.86	198	1.76	38	0.34	674	6.01	1789	15.94	2796	24.92	7353	32.76
Duplication Factor	203	1.81	0	0.00	1	0.01	0	0.00	1	0.01	463	4.13	668	5.95	809	3.60
ELI Factor	511	4.55	16	0.14	0	0.00	0	0.00	281	2.50	611	5.45	1419	12.65	2127	9.48
Geographic Factor	48	0.43	0	0.00	0	0.00	0	0.00	8	0.07	19	0.17	75	0.67	157	0.70
M Value Factor	2463	21.95	17	0.15	362	3.23	15	0.13	0	0.00	1750	15.60	4607	41.06	5708	25.43
% of POPS Factor	1332	11.87	11	0.10	147	1.31	29	0.26	137	1.22	0	0.00	1656	14.76	6288	28.02
Total	4557	40.61	141	1.26	708	6.31	82	0.73	1101	9.81	4632	41.28	11221	100	22442	100

Smoothed normalized basic weight estimates were computed for each month for each ELI in each index area by averaging the rebased basic weight estimates attached to the useable quotes across all PSUs and sample replicates. The objective here was to more evenly distribute a full sample estimate of daily expenditures among the sample quotes in an ELI within an elementary cell.

In single-ELI strata, except in the case of sub-sampling, this smoothing effectively equalized the weights of the sample quotes. In multi-ELI strata, this procedure had a similar effect among quotes for the same ELI, though there remained some variation due to differences in percent of POPS estimates among sample outlets. Though it was not entirely clear how modifications of this type would affect the value of the index series, it was expected that the procedure would be neutral with respect to aggregate index change and would likely dampen the sampling variability of full sample price relative and index series.

Price relatives were chained over time to produce monthly price indexes, and 1-, 2-, 6- and 12-month price change estimates for each elementary cell. Aggregate index and price change estimates were computed for each item stratum, major group and their combined series at the U.S. City Average level as well.

Sampling variance estimates were produced for each set of estimates using a stratified jackknife method, implemented in VPLX (Fay, 1990). The stratified jackknife variance computations in this application were based on a segmentation of the CPI sample into clusters within two separate Food strata, one Apparel stratum, and three Housing strata within self-representing index areas, and into clusters within one stratum for all items within non-self-representing index areas. The sample for the individual strata consisted of 31 clusters, with each cluster being the sample from one of 31 self-representing index areas. The sample for the sixth stratum consisted of seven clusters, each being the sample from one of seven non-self-representing index areas.

For the U.S. All Cities index, we constructed replicate indices $\hat{I}^u_{sm}$ ($u = t$ or t - k) for each of n_s clusters in each stratum s. We did this by deleting the index for cluster (m) in stratum (s), and aggregating the indices for the remaining clusters, adjusting or rescaling their weights, or relative importances, to produce an estimate for the full stratum. These were then aggregated with full sample estimates for the remaining strata. The index corresponding to replicate (sm) was:

$$\hat{I}^u_{(sm)} = \sum_{s'\in\{S-s\}} \sum_{m'\in\{M_{s'}\}} r^b_{s'm'} \hat{I}^u_{s'm'} + \sum_{m'\in\{M_s-m\}} r^b_{sm'} \hat{I}^u_{sm'}$$

where $\{M_s\text{-}m\}$ denotes the set of clusters for stratum s omitting cluster m, and $r^b_{sm'}$ is the rescaled relative importance for the stratum-cluster sm' in replicate (sm). Replicate estimates of k-month price change were derived by taking ratios of replicate indices:

$$\hat{I}^{t,t-k}_{(sm)} = \hat{I}^{t}_{(sm)} \Big/ \hat{I}^{t-k}_{(sm)}$$

The stratified jackknife estimator of the variance of $\hat{I}^{t,t-k}$ is given by

$$v_{sjk}\left(\hat{I}^{t,t-k}\right) = \sum_{s=1}^{S} \frac{(n_s - 1)}{n_s} \sum_{m=1}^{n_s} \left[\hat{I}^{t,t-k}_{(sm)} - \hat{I}^{t,t-k}\right]^2$$

where the first sum is over all strata in the sample.

3. Findings

At the elementary cell level, changes in quote level weights produced, in some instances remarkably different results from those found in production simulations, with differences ranging from -120% to +65%. However, these variations tended to cancel each other out at higher levels of item and area aggregation. Figure 1 depicts 1-month price change for selected food, apparel and housing item strata, at the U.S. City Average level, with and without quote weight smoothing for the 18 months of the study. Here we see that at the U.S. City Average level, there are generally only small differences between the 1-month price change series created using the smoothed and unsmoothed quote weight estimates. Within any item stratum cell they tracked each other fairly closely. This obtained for most item strata.

Table 2 summarizes differences in price change at the item stratum and major group level for index area and U.S. All Cities geographic levels. From this we see that the extreme differences observed at the elementary cell level did dampen in magnitude at the U.S. level. However, we also see that these differences produced positive averages and tended to accumulate with lag, at both the item stratum and major group levels. Differences in Food and Beverages were quite small compared to those for Housing C&S and Apparel.

Figures 2 to 5 depict 1- and 12-month price change and their standard errors for the study period for the U.S. City Average for each major group. Differences in 1-month price change were typically small. For Food and Beverages, there was one time interval, November to December 2002, in which the price change peaked differently for the two series. Though differences in 1-month changes were small, they appeared to accumulate in one direction over time, with 12-month changes for Housing C&S and Apparel appearing to diverge by the end of the study period.

For Food, standard error estimates are hardly distinguishable for these two series for any of the lags. For Housing, however, we see the anticipated dampening of variability for both 1- and 12-month price change in some of the months of the study. For Apparel, we see mixed results. The smoothing produced higher estimates of 12-month price change variability for two months of the study.

Figures 6 through 7 depict differences in 1- and 12-month price change plus and minus two-standard-error bands for the study period for the U.S. City Average for the aggregate index of the three major groups. We see here in the aggregate that for 12-month lags the differences in price change are above 0 for almost all the months of the study and the two-standard-error bands are above 0 for for the last three months, in what would appear to be an upward trend. This was investigated and determined to be principally attributable to large differences in HB02-Lodging Away from Home in New York suburbs A110 and A111, in Boston A103, and in X300 Medium-sized cities in the South. But large differences were not confined to these three index areas. In terms of weighted contribution to differences in the aggregate 12-month price change, this item stratum contributed over 40 percent of 12-month price change in each of the seven months for which this was estimated, outstripping other item strata by factors of 4 to 10. Other item strata with large positive weighted differences in many months were AC01-Women’s Outerwear, AC03-Women’s Suits and Separates, AF01-Infants’ and Toddlers’ Apparel, FK04-Other Fresh Fruits and FT06-Other Miscellaneous Foods (Figure 8.). On the average, food strata exhibited average price change differences an order of magnitude smaller than the other two major groups, though again this difference was positive.

Closer examination of the positive price difference phenomenon revealed that for weighted differences in 1-month price change between quote weighting methods, estimates for HB02 ranked first in aggregate contribution to positive differences for each month.

We examined the quote level behavior of the sample for HB02 to determine what systematically might be occurring. This investigation is ongoing, but preliminary analysis indicates that the combination of quotes with

Figure 1. 1-Month Price Change, with and without Basic Weight Smoothing for Selected Food, Apparel and Housing Strata, July 2002-December 2003

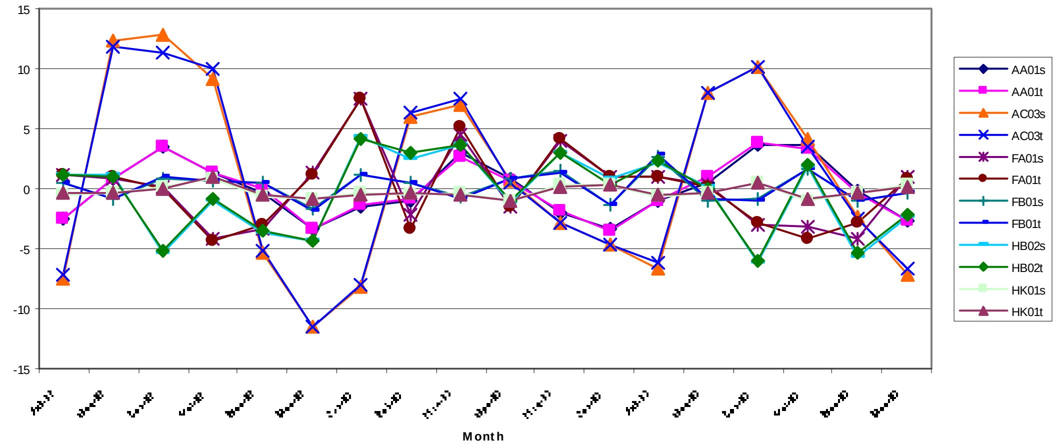


Figure 2. 1-Month Price Change, with and without Basic Weight Smoothing for Food and Beverages, Apparel, and Housing C&S, July 2002-December 2003

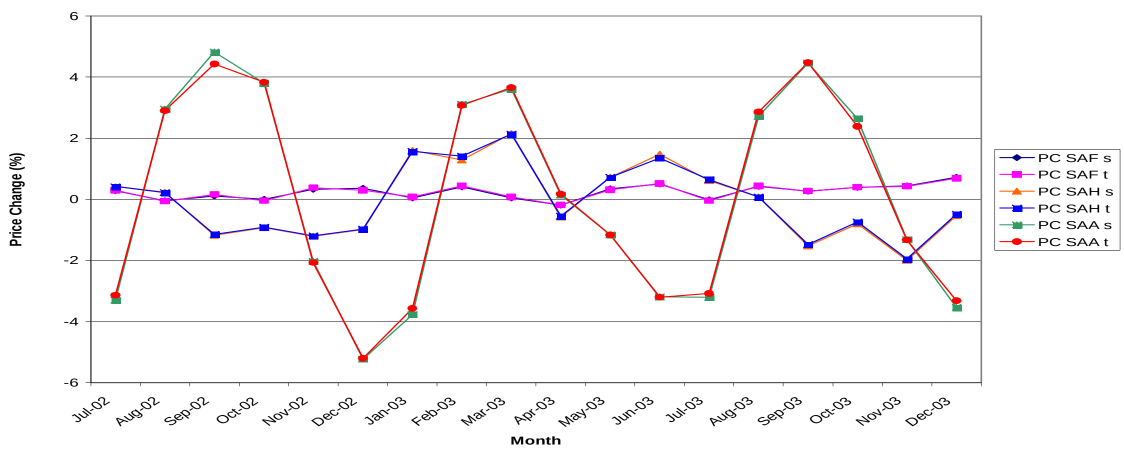


Figure 3. 1-Month Price Change SE, with and without Basic Weight Smoothing for Food and Beverages, Apparel, and Housing C&S, July 2002-December 2003

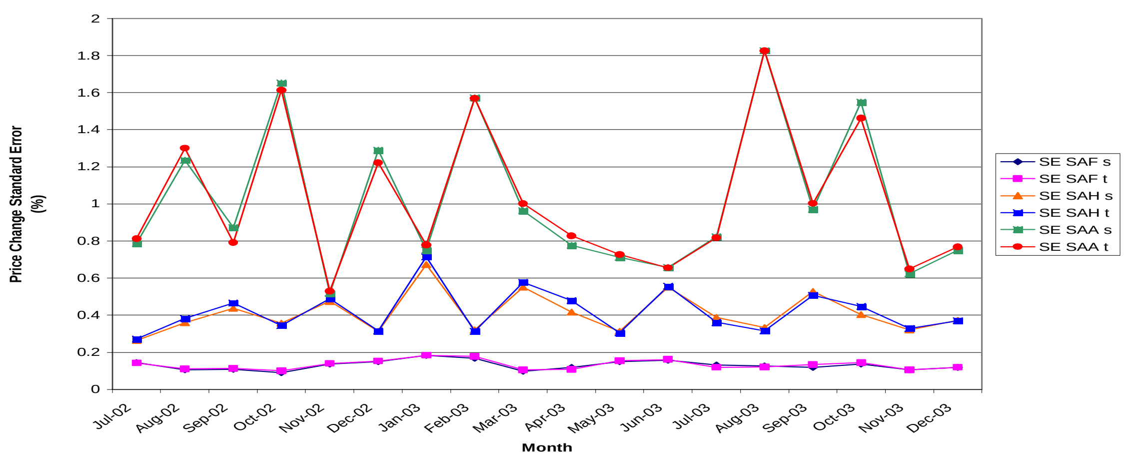


Figure 4. 12-Month Price Change, with and without Basic Weight Smoothing for Food and Beverages, Apparel, and Housing C&S, June-December 2003

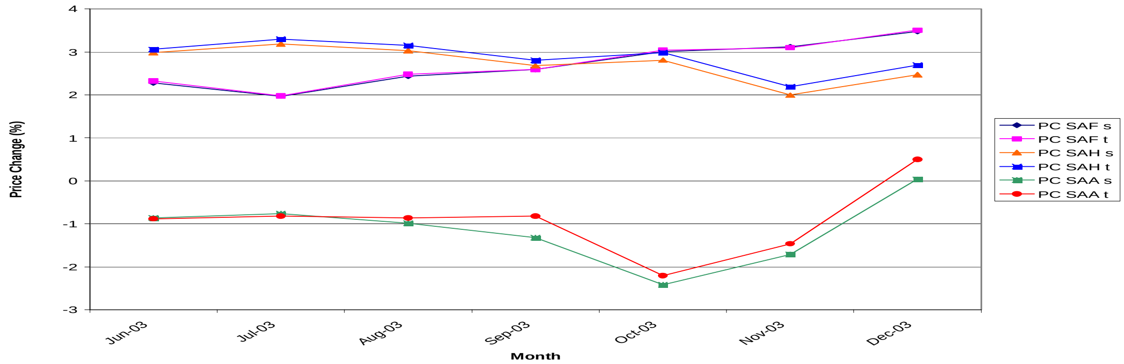


Figure 5. 12-Month Price Change SE, with and without Basic Weight Smoothing for Food and Beverages, Apparel, and Housing C&Sa, June-December 2003

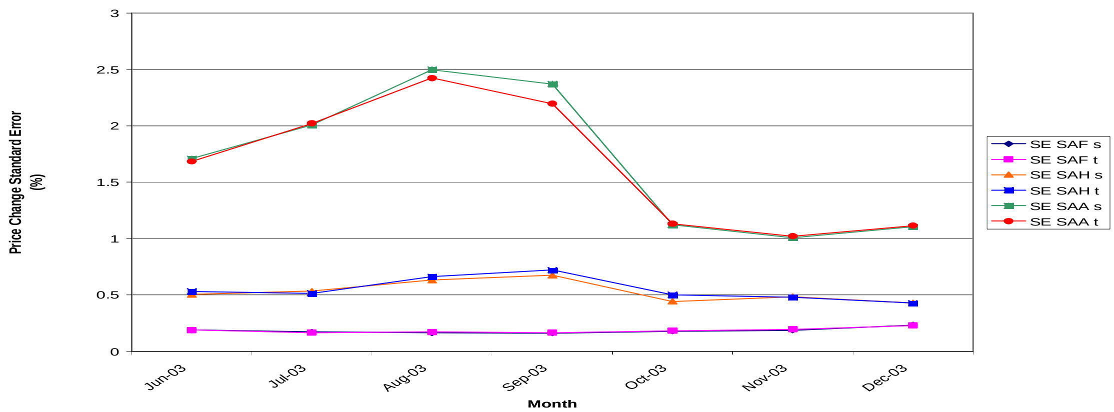


Figure 6. Difference in 1-Month Price Change +-2 Standard Errors, without vs with Basic Weight Smoothing for Food, Apparel and Housing, U.S. City Average, July 2002-December 2003

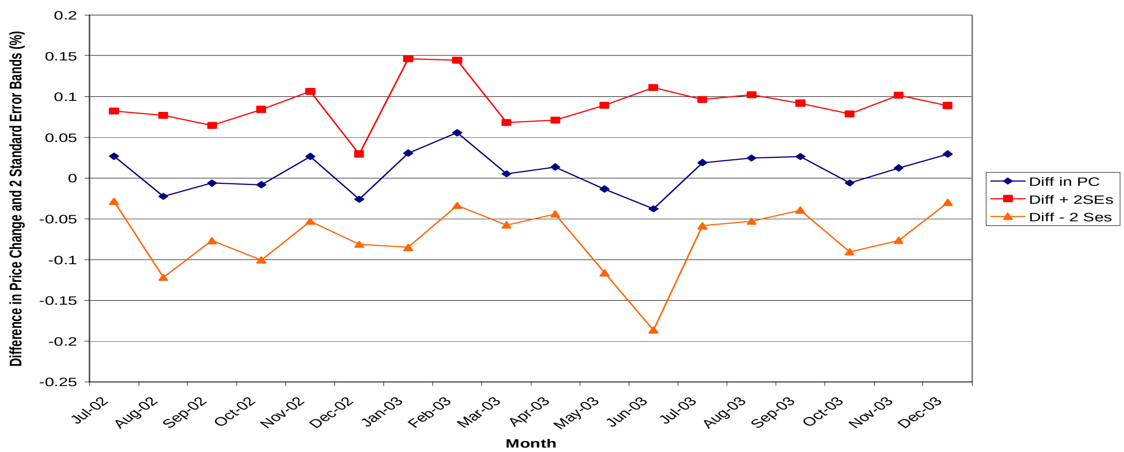


Figure 7. Difference in 12-Month Price Change +-2 Standard Errors, without vs with Basic Weight Smoothing for Food, Apparel and Housing, U.S. City Average, June- December 2003

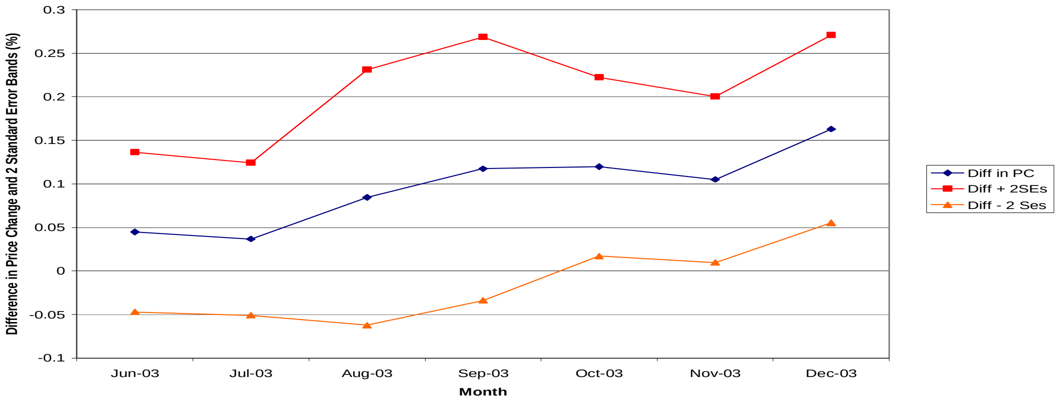
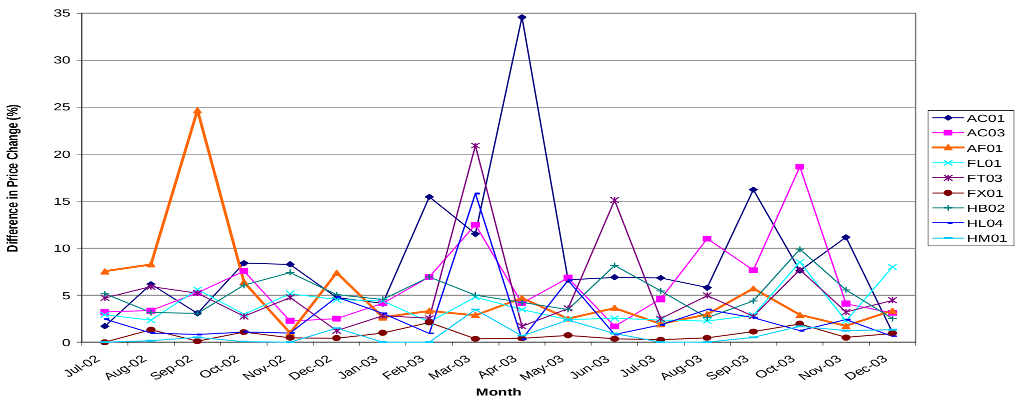


Figure 8. Largest Differences in 1-Month Price Change, with and without Basic Weight Smoothing for Selected Food, Apparel and Housing Strata, July 2002-July 2003



positive price change and smaller smoothed weights with those with negative price change and larger smoothed weights outstripped their counterparts in comparison of average price change in 10 of the 18 months of the study. Similar results were obtained for other item strata, though the differences in aggregate price change were not as large.

4. Conclusions

Smoothing of expenditure estimates across half samples is arguably justified from a full sample perspective. It tends to produce price change estimates with a dampened standard error, though for the item groups, variance estimation methodology and time interval of this study, these effects were small. For item groups with large positive quote level price changes, such as is observed in returns from sale and seasonal upswings in prices, this effective quote weight smoothing appears to produce lower aggregate price change estimates, which are not significantly different at the 1-month level, but which tend to accumulate over time and lag, producing significant differences for 12-month lags by the end of the study period. This phenomenon bears further investigation with respect to these and other CPI item groups and an expansion of the length of this study. It will be considered carefully before any weighting changes are recommended for publication index production.

5. Acknowledgments

The authors would like to thank Robert Cage, Alan Dorfman, and William Johnson for their careful reading of this paper and Janice Lent for helpful discussions regarding its findings. The authors would also like to acknowledge the contribution of Teague Ruder in constructing and providing the price relative computation software used in this analysis. The authors also thank William Johnson for his support on this project.

Table 2. Summary of Comparisons of Price Change for Food, Apparel, and Housing C&S Major Groups at Basic and Aggregate Levels by Lag for 200207-200312

Lag	# Obs	Mean Difference in Price Change, without vs with Smoothed Weights	Minimum Difference in Price Change, without vs with Smoothed Weights	Maximum Difference in Price Change, without vs with Smoothed Weights	Mean Price Change, without Smoothed Weights	Minimum Price Change, without Smoothed Weights	Maximum Price Change, without Smoothed Weights	Mean Price Change, with Smoothed Weights	Minimum Price Change, with Smoothed Weights	Maximum Price Change, with Smoothed Weights
Basic Cells (Item Stratum-Index Area)										
1	79344	0.0052	-120.3790	64.6812	0.5603	-81.7185	429.6386	0.5550	-81.7185	429.6386
2	74936	0.0041	-120.3790	64.6812	0.9986	-81.7185	429.6386	0.9945	-81.7185	429.6386
6	57304	0.0132	-63.0065	77.0130	1.8859	-81.1965	306.7157	1.8726	-81.1965	337.8217
12	30856	0.0164	-51.8317	54.9658	2.9333	-81.1965	334.5171	2.9168	-81.1965	334.8164
U.S. All City Averages for Item Strata										
1	2088	0.0028	-2.1388	3.6637	0.1787	-15.1226	16.5181	0.1759	-14.5835	17.1272
2	1972	0.0054	-3.1311	3.6153	0.4150	-22.9360	28.8118	0.4097	-22.6045	29.1217
6	1508	0.0233	-2.8212	4.3025	1.0548	-32.0019	54.6680	1.0315	-32.4434	54.8093
12	812	0.0487	-1.6971	3.2374	1.9866	-11.8048	38.2970	1.9380	-12.6395	39.2605
Index Area Averages for Food										
1	1116	-0.0014	-1.3055	1.5165	0.3042	-14.8937	16.5181	0.3056	-14.4949	17.1272
2	1054	-0.0015	-1.8479	1.6448	0.5731	-22.9360	26.9788	0.5746	-22.6045	26.9483
6	806	-0.0044	-2.1903	1.4916	1.8159	-32.0019	54.6444	1.8203	-32.4434	54.8093
12	434	0.0056	-1.6971	1.6717	3.5414	-9.8403	38.2970	3.5358	-10.5772	39.2605
Index Area Averages for Housing C&S										
1	612	0.0070	-2.1388	3.6637	0.0312	-15.1226	16.4586	0.0241	-14.5835	16.1415
2	578	0.0178	-2.7880	3.6153	0.0960	-22.0874	28.6139	0.0782	-21.7050	27.9747
6	442	0.0326	-2.8212	4.3025	0.6002	-27.0527	54.6680	0.5677	-26.9894	54.2660
12	238	0.0341	-1.6655	2.6762	1.1805	-10.3059	28.6897	1.1464	-10.1448	28.6897
Index Area Averages for Apparel										
1	360	0.0084	-1.9193	2.4316	0.0402	-11.5419	14.9647	0.0318	-11.5816	15.0014
2	340	0.0053	-3.1311	2.1891	0.4673	-18.6910	28.8118	0.4620	-18.8652	29.1217
6	260	0.0936	-2.2306	2.5599	-0.5319	-18.0927	19.8542	-0.6255	-20.6525	20.2668
12	140	0.2070	-1.5648	3.2374	-1.4629	-11.8048	18.5512	-1.6699	-12.6395	17.7144
U.S. All City Averages for Food										
1	18	0.0003	-0.0767	0.0543	0.0312	-0.1810	0.6907	0.2496	-0.1780	0.7233
2	17	0.0020	-0.0457	0.0578	0.0960	-0.1071	1.1279	0.4697	-0.1200	1.1631
6	13	0.0058	-0.0405	0.0626	0.6002	0.8274	2.2197	1.3507	0.8297	2.2533
12	7	0.0199	-0.0154	0.0432	1.1805	1.9739	3.5040	2.6964	1.9712	3.4739
U.S. All City Averages for Housing C&S										
1	18	0.0165	-0.1227	0.1215	0.0402	-1.9505	2.1379	-0.0619	-1.9917	2.1611
2	17	0.0325	-0.1166	0.1152	0.4673	-2.6690	3.5821	-0.1185	-2.7776	3.4815
6	13	0.0769	-0.0767	0.2236	-0.5319	-3.8684	6.8269	1.0086	-4.0920	6.8423
12	7	0.1458	0.0711	0.2241	-1.4629	2.1897	3.2958	2.7362	1.9937	3.1878
U.S. All City Averages for Apparel										
1	18	0.0101	-0.3872	0.2214	0.3042	-5.2001	4.4778	0.0893	-5.2217	4.8206
2	17	-0.0033	-0.4486	0.2655	0.5731	-8.5820	8.4374	0.6355	-8.7862	8.8039
6	13	0.0955	-0.2345	0.3208	1.8159	-4.1740	2.0551	-0.4921	-4.3741	2.0429
12	7	0.2099	-0.0608	0.5005	3.5414	-2.2047	0.4975	-1.1478	-2.4205	0.0375

6. References

Bureau of Labor Statistics (2003), *BLS Handbook of Methods*, Washington. DC, http://www.bls.gov/opub/hom/home.htm

Bureau of Labor Statistics, (1997), "The Experimental CPI using Geometric Means (CPI-U-XG)," April 10, 1997 (Washington: Bureau of Labor Statistics).

Fay, R. E., (1990), "VPLX: Variance Estimates for Complex Samples," *Proceedings of the Survey Research Methods Section*, American Statistical Association, pp. 266-271.