Evaluating a Calibration Weighting Scheme for Elementary Indexes for Commodities and Services in the U.S. Consumer Price Index

Sylvia G. Leaver, Robert A. Cage, and Darin T. Solk
U.S. Bureau of Labor Statistics, Washington, D.C. 20212

Abstract

The estimate of price change for an elementary cell of the U.S. CPI is a weighted average of quote-level price changes, where the quote-level weight is a function of inverse selection probabilities at each stage of selection. This paper evaluates, by means of retrospective estimation, two alternative weighting schemes employing calibration. The theory behind this calibration approach is that the sum of the weights of each “usable” quote in each elementary cell should yield an estimate of the expenditure for the elementary cell.

The CPI produces two distinct estimates of elementary item-area expenditure: (i) that derived from the sum of Commodities and Services (C&S) expenditures reported in the Telephone Point of Purchase Survey (TPOPS), and (ii) that derived from Consumer Expenditure Survey (CEX) biennial data. The first, more limited, weighting scheme ratio-adjusts TPOPS expenditure estimates in each stratum-PSU-replicate panel to be proportionate to their sub-stratum component, or Entry-Level Item (ELI), selection probabilities. The second calibration adjustment forces quote weights in any given month to sum to biennial expenditure estimates from the CEX. In Section 1 we provide further background on the CPI and price relative estimation. In Section 2 we describe the two calibration formulas and present their rationales. In Section 3 we give our findings from a retrospective study, and compare the behaviors of test calibration index series with the CPI production series. In Section 4 we present our conclusions and discuss further research directions in this area.

Keywords: Consumer expenditures, Simulation, Ratio adjustment

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

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 by 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 (SR) areas, such as New York, which were selected with certainty, and non self-representing (NSR) areas, the samples of which comprise two or more primary sampling units (PSUs) selected according to a probability sample. The U.S. City Average CPI is a weighted average of 38 index area CPIs—31 from SR and 7 from NSR areas. For the purposes of variance estimation and operational manageability, the sample for each self-representing PSU is segmented into two or more subsets, called sample replicates, or historically, half-samples (HS). 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.

The CPI is estimated for items grouped into 211 strata for each index area, although not all such indexes are published every month. Each item stratum is composed of one or more narrowly defined groupings called 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 selection probabilities are derived from expenditures reported in the two most recent years of the CEX, summarized at the Census regional level. ELI selections are independently drawn for each HS within each PSU.

Sample frames and weights used in outlet selection are derived from 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-HS.

In outlet selection, outlets are selected via systematic pps from frames for each PSU-HS for POPS categories corresponding to ELIs selected in item sampling. Selected items, termed quotes, are then priced in sample outlets on a monthly, bimonthly, or seasonal basis. We note here that most samples in the CPI are priced bimonthly instead of monthly. For bimonthly priced items in NSR areas, roughly one-half of the PSUs are priced on an even-month cycle, and the other half on an odd-month cycle.

The CPI is constructed in two stages. In the first—or elementary cell—stage, the price index for an item-area-cycle is updated every one or two months via a price relative, a function of sample quote-level price changes. In the second stage the elementary cell indexes are aggregated, by means of a weighted average, to produce price indexes for higher level item and geographic groupings. We will focus on the first stage in our exposition. Let X_{iac}^{t} denote the index at time *t*, in item stratum *i*, area *a*, and cycle *c*, relative to time period 0. Then $X_{iac}^{t} = R_{iac}^{t,t-1} X_{iac}^{t-1}$ where $R_{iac}^{t,t-1}$ denotes the on-cycle price relative between times *t* and *t*-1 . Since 1999, price relatives for most commodities and services have been computed using the following weighted geometric average (BLS, 1997): $R_{iac}^{t,t-1} = \prod_{k \in S_{iac,t}} \left(P_{k,t} / P_{k,t-1} \right)^{W'_{k,t}}$.

The price relative formula for a comparatively smaller remaining set of commodities and services is a modified weighted Laspeyres formula:

$$R_{iac}^{t,t-1} = \frac{\sum_{k \in S_{iac,t}} \frac{W_{k,t} P_{k,t}}{P_{k,b}}}{\sum_{k \in S_{iac,t}} \frac{W_{k,t} P_{k,t-1}}{P_{k,b}}}.$$ Here

$S_{iac,,t}$ represents the sample for item *i* in area *a* and cycle *c* at time *t* which are eligible for price relative calculation (PRC);
P_k represents the price for quote *k*;
$P_{k,b}$ represents a TPOPS reference base period price for quote *k*, used to convert TPOPS expenditures to quantities in the Laspeyres formula;
$W_{k,t}$ represents the quote-level expenditure weight of sample observation *k* at time *t*; and
$W'_{k,t}$ represents the same for item *k*, normalized to the same sample rotation base for all quotes in the item-index area and expressed in share form:

$$W'_{k,t} = \frac{\dfrac{W_{k,t}}{X_{iac,b} / X_{iac,Dec\,1997}}}{\sum_{j \in S_{iac,t}} \dfrac{W_{j,t}}{X_{iac,b} / X_{iac,Dec\,1997}}}$$

This paper will focus on the quote-level expenditure weights, whose structure is given below:

$$W_{k,t} = \frac{\alpha \cdot E \cdot f \cdot g \cdot \gamma}{\beta \cdot M_t}$$, where

α is the percent of sales of the ELI to the total sales of the POPS category in the sample outlet, collected in the CPI pricing survey at initial item selection,
E is the basic weight, an estimate of total daily expenditures for the POPS category, for the CPI urban population in the PSU-HS, derived from the POPS survey. We note here that 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 in every PSU in each NSR index area nor in all HSs in the three largest SR PSUs.
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 population of the PSU between index revisions.
M is the number of usable quotes for the ELI-PSU replicate.
γ_k is the share of *n* stratum selections that were selected for the ELI in the PSU-HS.
β is the probability of selection of the ELI within the stratum in the PSU-HS, which is, in most cases, the proportion of the expenditure for the ELI of the total expenditure for its item stratum, reported in the CEX survey in its Census region.

Under certain assumptions (Cage, 2005; BLS 2004) the formula reduces to:

$$W_{k,t} = \frac{\mathrm{E}_{iacsh,b}}{\mathbf{n}_{iacsh,t}}$$

where $E_{iacsh,b}$ is the expenditure estimate for the item stratum *i* in index area *a*, cycle *c*, PSU s and half-sample *h* during base period *b and* $\mathbf{n}_{iacsh,t}$ is the sample size for the same at time *t*. That is, the weight of each quote in a given month *t* is equivalent to an expenditure estimate on the corresponding item-area-cycle-PSU-HS at base period *b*, divided by the number of active quotes (sample size) available for that same area in month t. Note that if all sources of variability are removed from this construction, the resulting weights across the set of all quotes within an elementary item-area-cycle cell would be nearly equal, varying only by the differences in expenditure estimates and sample size among constituent PSU-HSs.

The structure of the CPI's elementary indexes, with respect to both the geographic component (index area) and the consumer goods and services component (item stratum), defines the target expenditure level that quote weights are designed to represent. That is, generally, the sum of the final weight over all quotes belonging to the active sample of quotes for elementary item i in PSU s and HS h in elementary area a should yield an estimate of the aggregate spending on the item-area-cycle-PSU-HS in base period b.

We note here that the three surveys from which component quote-weight variables are calculated operate at various levels of geographic and consumer good/service classification. Ultimately, quote weights are derived from the product of these different estimates. Use of multiple survey sources, compounded by the interchange of varying geographic-item classifications, may produce undesirable variability and bias in quote weight estimation.

Quote weights are more likely to vary within an elementary item-area cell if the item stratum is composed of multiple ELIs, the index area is composed of multiple PSUs, the POPS category is composed of multiple ELIs, there is more than one POPS category for the item stratum, or sub-sampling is performed in one or more outlets

Variability in the estimates of average daily POPS category expenditure across PSU-HSs also contributes to variability in quote weights at the elementary level. We note here that in single-ELI single-PSU elementary cells where sub-sampling does not occur, the only source of variation in quote weights is the difference in the estimate of POPS category daily expenditure in the two HSs.

If little quote weight variation were present in an elementary cell, then the ratio of the largest quote weight to the smallest quote weight would be near 1. In reality, this ratio varies. In December 2004, the mean ratio of high to low weight was 8.4. For multiple-PSU index areas, the mean ratio was 21.5, and for single-PSU index areas, it was 3.6. For multiple-ELI item strata, the mean ratio was 14.1, and for single-ELI item strata, it was 5.7.

The greatest degree of variation in relative quote weights occurs in multiple-PSU (i.e., NSR), multiple-ELI elementary cells. Experimental evidence suggests the CPI estimate of price change at the All Items U.S. City Average level would be slightly lower if the quote weights in the elementary cells were equal. Between 2000 and 2004, the average 12-month price change of the CPI (C&S items only) was 2.26 percent. With uniform quote weights at each elementary cell, the average 12-month price change dipped to 2.21 percent—smaller by 0.05 percentage points. This suggests that removing certain sources of variability from quote weight estimates would produce lower estimates of aggregate price change.

A similar finding held in another investigation of quote weight adjustment procedures as well. Leaver and Solk (2004) studied the effect on the level and variability of price change estimates produced by smoothing of POPS expenditure estimates across PSU-HSs within index areas. For item groups with large, positive quote-level price changes, such as is observed in returns from sales and seasonal upswings in prices, and for index areas with multiple PSUs, this quote-weight smoothing produced lower aggregate price change estimates than those derived from production values,. These differences were not significant at the 1-month level, but accumulated over time and lag, becoming significant for 12-month lags by the end of the study period.

Some improvements and efficiencies in quote weights might be realized by changing the elementary item-area structure of the CPI or by redefining the sampling unit of the component surveys (e.g., collapsing POPS categories in TPOPS or eliminating multiple-ELI item strata). However, changes to the CPI market basket structure affect not only quote weight estimation, but also all aspects of CPI estimation, from price relative calculation and imputation, to aggregation weight computation, and to data collection procedures and resource management. From a program perspective, significant revisions to the market basket structure are relatively time-consuming and resource-intensive, and thus are infrequent. The CPI area sample is typically revised every 10 years. The CPI item structure was revised with the 1978 revision and again (with minor changes) with the 1987 revision. There were additional significant changes with the 1998 revision, and none since.

2. Calibration formulas

An alternative approach to reducing undue variability of quote weights is calibration. We recall that currently, the CPI has available two distinct estimates of elementary item-area expenditure: (i) that derived by the sum of C&S quote weights, and (ii) that derived directly from CEX biennial data. Further evidence also suggests that between-PSU price-change variation can be significant for some item classes (Shoemaker, 1999 and 2001, and Leaver-Larson, 2003.) It thus seemed reasonable to consider controlling for PSU market share of item expenditures within an elementary item-area cell. We examined two different methods to resolve inconsistencies between the two estimates.

The first method calibrated TPOPS expenditures reflected in E, the quote basic weight term, so that the within stratum-PSU-HS share of TPOPS-based expenditures for each member TPOPS category equaled the same share of ELI expenditures derived from the CEX for those ELIs belonging to the TPOPS category:

$$\widetilde{W}_{k,t} = W_{k,t} \cdot \frac{\sum_{POPS,\, l \in (iacsh)} E_{POPS,l}}{E_{POPS,k}} \cdot \frac{\sum_{eli,e \in POPS,l(iacs)} E_{CE,e,r}}{\sum_{eli\, e \in (iacs)} E_{CE,e,r}}$$

The first term in the ratio adjustment is the inverse of the within-stratum TPOPS-based expenditure share for the TPOPS category to which quote k belongs. The second term in the ratio adjustment is the within-stratum CEX-based expenditure share for the same. This ratio adjustment prevents the exaggeration of quote weights, which occurs when a large basic weight value representing a disproportionately larger within-stratum TPOPS expenditure value is divided by a very small ELI selection probability. There are 40 priced item strata with multiple TPOPS categories, representing between 18 percent and 19 percent of C&S expenditures, that are eligible for this ratio adjustment.

The second method sought to eliminate inconsistencies between the two estimates by calibrating quote weights in each month to force their sum to equal the elementary biennial expenditure estimates from CEX, which are used to construct CPI-U aggregation weights in that month.

Specifically, the weight of each quote k belonging to an ELI e in an item stratum-area-cycle-PSU- *(iacs)* could be set equal to the estimate of aggregate annualized expenditure on the ELI e in the PSU, divided by the corresponding number of selected quotes per ELI-PSU. If controlling for market share at the ELI level were deemed unwarranted, the weight of each quote k could be set equal to aggregate annualized expenditures on the item i in the PSU.

For geographic calibration, the weight of each quote was recalculated as its corresponding item-area-cycle-PSU annualized expenditure from the 2001-2002 CPI-U expenditure reference period, divided by the number of "active" quotes in the item-area-cycle-PSU:

$$\hat{W}_{k,t} = \frac{E_{iacs}}{n_{iacs,t}}, \text{ where}$$

E_{iacs} = the CEX-based annualized expenditure estimate for the corresponding item-area-cycle-PSU for the active CPI-U expenditure reference period b. For January 2004 to December 2005, b={2001, 2002}, and

$n_{iacs,t}$ =the number of quotes in the corresponding item-area-cycle-PSU that were used in price relative calculation in month t.

For SR index areas, the sample PSU is the entire index area, so each item-area-cycle-PSU expenditure E was set equal to the composite estimated annualized item-area expenditure derived from CEX data, from the 2001-2002 biennial time period. NSR areas are composed of multiple PSUs, so the estimation of item-area-PSU expenditure for them was a two-step procedure. First, item-area-cycle expenditures were estimated by multiplying the composite estimated item-area expenditure derived from CEX data by the ratio of item-area-cycle to item-area preliminary expenditure estimates:[1]

$$\mathrm{E}_{iac,b} = E_{ia,b} \times \frac{\sum_{s \in iac} \widehat{E}_{ias,b}}{\sum_{s \in ia} \widehat{E}_{ias,b}}$$

If one or both cycles in an area had missing expenditures for an item (no expenditure reports in CEX for the 24 months of 2001-2002), the ratio of 1990 item-area-cycle to item-area population was used instead of the ratio of

[1] Preliminary estimates were derived from annualized 2001-2002 CEX data prior to composite-estimation (underlying sample sizes of item-area-cycle roughly half that of item-area.)

expenditures.[2] In the study period, this occurred 3 times.[3] Next, the item-area-cycle expenditures, $E_{iac,b}$, were allocated into each PSU belonging to the area-cycle based upon the 1990 population PSU-to-area-cycle shares:[4]

$$E_{iacs,b} = E_{iac,b} \times \frac{POP_{s,1990}}{\sum_{s \in c} POP_{s,1990}}$$

If one or more PSUs in an area-cycle had no active or usable quotes in a given month t (i.e., if $n_{k \in S_{ias,t}} = 0$), then the item expenditure in that PSU was ratio-allocated to the remaining quotes in the area-cycle such that the sum of the quote weights within the area-cycle equaled $E_{iac,b}$ Similarly, if one area-cycle in an area had no active quotes, its item expenditure was allocated to the other cycle such that the sum of the quote weights by item-area equaled $E_{iac,b}$.

The perceived advantages for considering Method 2 calibration were numerous. First, the sum of quote weights by item-area-cycle yields an estimate of item-area-cycle expenditure explicit in the estimation of CPI-U aggregation weights. That is, the sum over all quotes in each elementary cell would yield the item-area biennial expenditure used to derive CPI-U aggregation. Weights are kept constant for 24 months, then updated concurrently with expenditure weight update in the CPI-U. Second, the estimate is derived from a single-source (CEX), as opposed to the current approach, which uses three sources (C&S, TPOPS, and CEX), and thus would eliminate the disagreement between the implicit time period of estimates for α, β, and basic weights. Use of this weight also eliminates the need and cost of collecting α, and the use of *f*, and obviates the need to rebase basic weights with price indexes to derive geomean weights, as expenditures are already on a common base.

Method 2 was also seen as a potential improvement over current weights in that both the implicit price and the quantity component of the expenditure estimated would be from the same period, so there would be no quantity bias in favor of more recently rotated items.

Operationally, the geographic calibration method is easy to implement from a systems perspective. Weights for the entire sample are calibrated for the same time period, which would eliminate potential negative consequences on weights in the form of seasonality variation caused by varying rotation times in TPOPS.

Additionally, this calibration eliminates the need for imputation at the ELI-outlet level. Currently, if sub-sampling is performed at an outlet, the weight of unsampled quotes is absorbed by sub-sampled quotes via *f*. Under this calibration approach, the weight of an unsampled quote is absorbed by all remaining quotes in the item-PSU. This has the potential advantage of reducing the weight of outliers—and possible price change outliers—absorbed by all quotes in the item-PSU, which is a larger set than ELI-PSU in multiple-ELI cells.

3. Findings

We first examined the effect of the weighting adjustments on elementary cells, using CPI quote-level data collected for the 24 month period, January 2004-December 2005. Tables 1 and 2 present summaries of quote weight distributions and price change, respectively, by lag, in elementary cells, for the three weighting treatments examined: a production simulation with no quote weight adjustments; Method 1, the TPOPS ratio-adjustment method and Method 2, the geographic calibration. All tables and figures are based on price relative calculation for all index areas for the 180 priced C&S strata.

From Table 1 we see that both experimental weighting methods produced fewer extreme-valued weights than were found in production. Not surprisingly, the effect of Method 2 was

[2] Calibration on expenditure share is preferred over calibration on population share (population share calibration assumes spending, by item, is identical across PSUs within the same area-cycle). Currently, item-area-cycle aggregation weights for the CPI-U are estimated using population share split factors instead of expenditure share split factors, so use of expenditure share here could be deemed as an improvement.

[3] The item-area-cycles split by population share instead of expenditure share were: Used Cars and Trucks in area=D400, and Other Motor Fuels in area D200 and D400.

[4] Generally, there is insufficient sample size at the item-PSU level, even when using 24 months of data, to yield reliable expenditure estimates (item-psu expenditures were missing 5% of the time for 2001-2002, with the median number of expenditure reports per item-psu equal to 25.) Thus population shares were used as an alternative.

more pronounced than that of Method 1, as Method 1 adjusts weights for only a subset of priced strata. For single-PSU areas, Method 2 effectively equalized quote weights and dampened the mean high-to-low weight ratio.

With respect to elementary cell-level price change, our findings were consistent with those from the 2004 Leaver-Solk study. Method 1 produced, on average. price change estimates that were only slightly different from production. Method 2 produced a more pronounced effect, with price change estimates that averaged lower than production values, and differences that accumulated with lag. Figure 1 gives the frequency distribution of elementary 24-month price change for multi-PSU Areas, for Method 2, overlaid on the same for production estimates. It demonstrates the degree to which the weighting adjustment reigned in both tails of the price change distribution and shifted the mean and median to the left.

These findings were echoed in our comparisons at the aggregate level, which are summarized in Table 3. Again, the two weighting adjustments tended to produce lower price change estimates. For 1-month price change, Method 1 differed on the average only slightly from the production treatment, while Method 2 was a bit lower. As was the case at the elementary cell level, differences between production estimates accumulated more for Method 2 than for Method 1 for longer lags, with differences in 12- and 24-month price change between the methods quite apparent by the end of the study period. Differences in 24-month price change between the production simulation and Methods 1 and 2 were .0032 and .1721 percentage points, respectively.

Standard errors for and differences in price change were computed using a stratified jackknife method implemented with VPLX (Fay, 1990.) Figure 2 graphs differences in cumulative price change between production estimates and Methods 1 and 2, with their 2- standard error bands for the study period. Method 2 differences were significantly different from 0 for 11 of the 24 months of the study.

Figure 2 also depicts the seasonal behavior of these differences, which peaked for both methods in April-May and again in October-November, the points at which new quotes associated with semiannual TPOPS sample rotations are introduced. It is in the months of new sample-quote introduction that positive differences in price change estimates were greatest and most numerous.

We explored these effects further, concentrating on differences in 1-month price change in October of each year for Method 2, and examining the weighted contribution of each item-area cell to the aggregate difference. For October 2004, Women's Suits, Motor Fuels, Clocks, Lodging While Out of Town, and New Vehicles were principal contributors to differences. Lodging While Out of Town, which is heavily weighted, and both highly seasonal and variable, dominated in October 2005.

Figure 3 presents 1- and 12-month price change standard errors for the three series. Variances for price change averaged, in general, lower for the calibration methods than those for production. At the U.S. City Average, All Items less Rent and OER level, they averaged less than 1 percent and about 5 percent lower for Methods 1 and 2, respectively, for 1-month change. For 12-month price change, the variance improvement was negligible for Method 1, but for Method 2 averaged over 10 percent, representing about a 5 percent decrease in standard error.

4. Conclusions

Expenditure-based calibration adjustments of quote level weights for price index estimators, such as those described for Methods 1 and 2, can be defended in terms of making quote weight factors more congruent, as in Method 1, or more timely, as in Method 2. From this retrospective experiment, each method reined in extreme-valued quote weights for affected strata, with differences apparent in resulting aggregates. For each method, the calibration adjustment produced, in the aggregate, lower price change estimates than those from production, though these differences were very small for Method 1. For Method 2, differences were smaller at the 1-month level, but accumulated over time, lag, and item aggregation, so as to be noticeable for All C&S for 12- and 24-month lags by the end of the study period. Certain item strata—in particular Lodging While Out of Town, Women's Apparel, and New Cars—were principal contributors to these consistent differences, which for Method 2 peaked for months in which new sample quotes were introduced. Each method produced price change estimates with smaller variances, though these effects were slight for Method 1.

These phenomena bear further investigation and extending the length of the study period seems advisable. In particular, given the relative ease of implementing Method 2, expected savings in data collection and respondent burden, and decreased sampling variability, this method bears further scrutiny and evaluation. Our findings will be considered carefully before any weighting recalibration is recommended for index production.

5. Acknowledgments

The authors would like to thank Alan Dorfman for his careful reading of this paper and helpful suggestions. The authors also thank John Layng, Robert Eddy and William Johnson for their support of this project.

6. References

Bureau of Labor Statistics (2003), *BLS Handbook of Methods*, Washington, D.C. 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)

Cage, R. A., et al. (2005), "Recommendations for Improving C&S Quote Weights," memorandum from the Quote Weight Review Team to the CPI Revision Planning Group, September 16, 2005, internal memorandum.

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

Leaver, Sylvia G., and Larson, William E., (2003), "Estimating Components of Variance for the U.S. Consumer Price Index," *Proceedings of the Survey Research Methods Section*, American Statistical Association.

Leaver, Sylvia G., and Solk, Darin T., (2004), "Estimating Sampling Weights for the U.S. Consumer Price Index," *Proceedings of the Survey Research Methods Section*, American Statistical Association.

Shoemaker, O. J., and Johnson, W. H. (1999) "Estimation of Variance Components for the U.S. Consumer Price Index", *Proceedings of the Section on Government Statistics, American Statistical Association*, 298-303.

Shoemaker, O. J. (2001) "Estimation of Variance Components for the U.S. Consumer Price Index: A Comparative Study", *Proceedings of the Joint Statistical Meetings, American Statistical Association.*

Table 1. Summary of Quote Weight Distribution in Elementary Cells, Production vs Methods 1 and 2

	Mean Low Weight Share	Mean High Weight Share	Mean Ratio High / Low	RMSD from Equal Weights	Mean Share
All C&S Priced Elementary Cells					
Production Weighted	0.086	0.180	7.890	0.036	0.119
Method 1	0.089	0.175	7.134	0.033	
Method 2	0.115	0.131	1.537	0.006	
AREA-CYCLE TYPE:					
Single-PSU					
Production Weighted	0.105	0.189	3.419	0.035	0.136
Method 1	0.107	0.184	3.045	0.032	
Method 2	0.136	0.136	1.000	0.000	
Multiple-PSU					
Production Weighted	0.046	0.159	17.575	0.038	0.079
Method 1	0.047	0.155	16.132	0.036	
Method 2	0.068	0.119	2.718	0.018	
ITEM TYPE:					
ELI = ITEM = POPSCAT					
Production Weighted	0.095	0.160	4.924	0.027	0.118
Method 1	0.095	0.160	4.924	0.027	
Method 2	0.114	0.130	1.564	0.006	
ELI ≠ ITEM ≠ POPSCAT					
Production Weighted	0.073	0.210	12.381	0.049	0.120
Method 1	0.079	0.198	10.480	0.043	
Method 2	0.115	0.132	1.496	0.006	

Table 2. Summary of Stratum Level Price Change at Index Area Level

Item Group	Lag	# Obs	Mean PC, Produc- tion	Mean PC Meth 1	Mean PC Meth 2	Mean Diff, Prod-Meth 1	Mean Diff, Prod-Meth 2	Min Diff, Prod-Meth 1	Min Diff, Prod-Meth 2	Max Diff, Prod-Meth 1	Max Diff, Prod-Meth 2
All Priced C&S Items	1	4320	0.18546	0.18541	0.1792	0.00006	0.00622	-1.09818	-2.02012	1.11010	3.05286
All Priced C&S Items	2	4140	0.39980	0.39922	0.3872	0.00059	0.01263	-1.78672	-3.62939	1.60295	4.15009
All Priced C&S Items	6	3420	1.09014	1.09215	1.0688	-0.00201	0.02135	-2.31869	-4.68987	2.07415	4.56872
All Priced C&S Items	12	2340	2.11041	2.10753	2.0672	0.00288	0.04320	-2.23810	-4.61757	2.63651	4.71843
All Priced C&S Items	24	180	4.54040	4.54176	4.3832	-0.00137	0.15718	-2.29870	-2.54647	3.06547	3.84159

Figure 1. Frequency Distribution of Elementary 24-Month Price Change, for Muliple-PSU Areas, Production vs Method 2

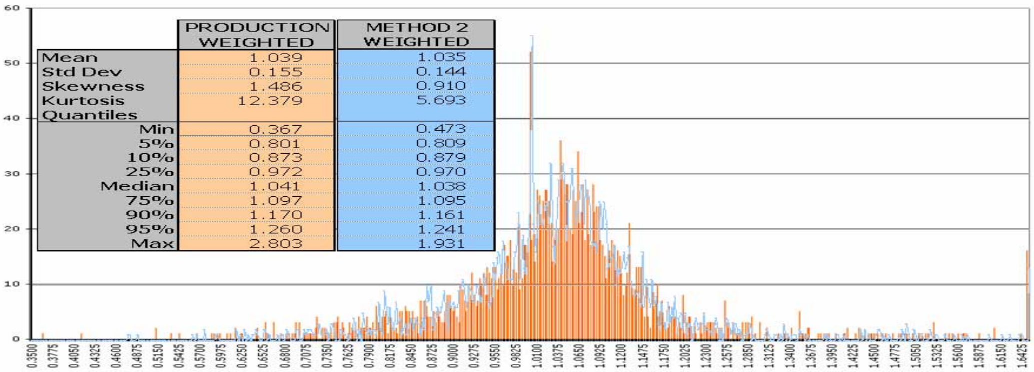

Table 3. Summary of Price Change for All C&S Priced Items at U.S. City Average Level											
Item Group	Lag	# Obs	Mean PC, Produc-tion	Mean PC Meth 1	Mean PC Meth 2	Mean Diff, Prod-Meth 1	Mean Diff, Prod-Meth 2	Min Diff, Prod-Meth 1	Min Diff, Prod-Meth 2	Max Diff, Prod-Meth 1	Max Diff, Prod-Meth 2
All Priced C&S Items	1	24	0.30353	0.30339	0.2967	0.00014	0.00680	-0.00915	-0.07094	0.01223	0.07795
All Priced C&S Items	2	23	0.63944	0.63912	0.6260	0.00032	0.01346	-0.01343	-0.13730	0.02028	0.14366
All Priced C&S Items	6	19	1.99958	2.00287	1.9889	-0.00329	0.01069	-0.02659	-0.08169	0.02374	0.14404
All Priced C&S Items	12	13	3.74370	3.74454	3.7108	-0.00084	0.03295	-0.01188	-0.04837	0.01494	0.12909
All Priced C&S Items	24	1	7.49554	7.49239	7.3234	0.00316	0.17210	0.00316	0.17210	0.00316	0.17210

Figure 2. Difference in Cumulative Price Change, Production vs Methods 1 & 2, U.S. City Average, January 2004-December 2005

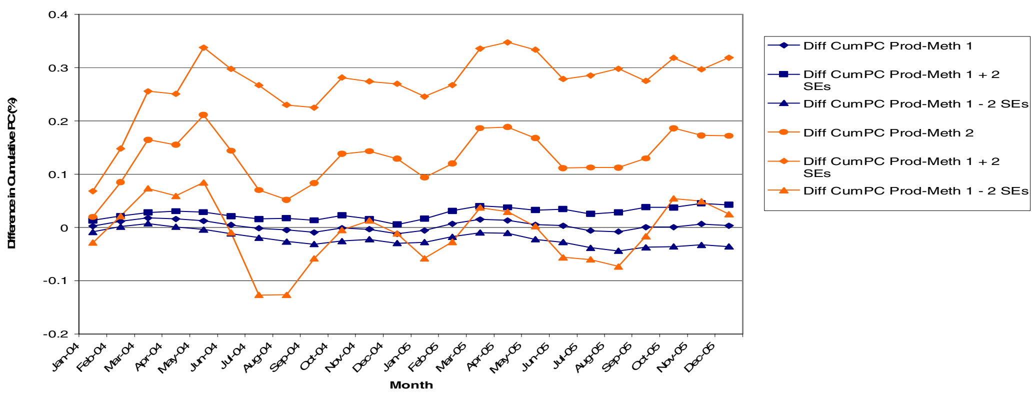

Figure 3. 1- and 12-Month Price Change Standard Errors, by Weighting Method, U.S. City Average, January 2004-December 2005