ESTIMATING UNCONDITIONAL VARIANCES FOR THE U.S. CONSUMER PRICE INDEX FOR 1978-1986

Sylvia G. Leaver, James E. Johnstone, Kenneth P. Archer, U.S. Bureau of Labor Statistics
Sylvia G. Leaver, U.S. Bureau of Labor Statistics, 600 E Street, N.W., Room 5217, Washington, D.C. 20212

KEY WORDS: replication, random groups, expenditure weights, linearization

In this paper we present final estimates of the sampling variance of the Consumer Price Index, and variance estimates of price change reflected by the index, for the years 1978 through 1986. These estimates are the final products of a three- phase estimation effort. In the first phase estimates of index and price change variances which are conditional on the December 1977 expenditure weights derived from the 1972-1974 Consumer Expenditure Surveys were produced. In the second phase estimates of the December 1977 expenditure weight variances were computed, and in the third phase these estimates were combined to produce estimates of unconditional variances for the index and price change. Variances of expenditure weights were estimated using balanced repeated replication. Conditional variances for indexes were estimated via replication using random groups, and variance estimates for price change were estimated via linearization, using index variance and expenditure variances. This paper presents estimates of index and 1-, 2-, 6-, and 12-month price change variance, which incorporate the sampling variance of the December 1977 expenditure weights, computed at the national level for All Items, seven major groups, and three housing subgroups, and at the regional level for All Items.

1. Introduction and Findings

This paper marks the first release of unconditional variance estimates for the Consumer Price Index. These estimates are for the time period January 1978 through December 1986. It follows preliminary estimates for the variance at the U.S. level of the index and for price change, which are conditional on the December 1977 expenditure weights, which were published in these proceedings in August 1990.

The behavior of unconditional variances of the CPI were seen to be little different from that observed for conditional variances. The variability in index and price change estimates which is attributable to variation in expenditure weight estimation is generally small, ranging from negligible for index variance estimates to between 6% and 20% for 12 month price change estimates for all items and major groups at the U.S. All Cities level. Unconditional variance of the CPI is an increasing function of time and value of the index, for the U.S. All Items index as well as for the indexes published at the U.S. level for major groups of commodities and services and housing. The variance of price change exhibits varying behavior over time. It is quite stable for the All Items index and most major groups but exhibits an upward trend and considerable variability for Apparel in particular. The methodology for estimating conditional and unconditional variances is given and sources of variability in variance estimates are discussed.

2. Background: The Consumer Price Index

For a full discussion of the CPI for 1978-1986, we refer the reader to the *BLS Handbook of Methods* (1988). A description of the features of the CPI which pertain to estimating variances is also given in Leaver (1990).

The CPI is estimated on a monthly basis for all consumer items for the total urban U.S., as well as at numerous other levels defined by the geographic area, group of items, and population (type of consumer). There are two populations for which the CPI is presently computed: urban wage earners and clerical workers (CPI-W) and all urban consumers (CPI-U). This report focuses on estimates of variance of the index and price change computed for the all urban consumers population.

We shall refer to the following item classifications:

(a) The *item stratum* is the most refined classification of commodities or services for which estimates of expenditures are computed and indexes are published. Examples are bread and college tuition.

(b) Item strata are grouped into *expenditure classes* (EC's). Examples include bakery products and educational expenses.

(c) Expenditure classes are combined into seven published *major groups*: food, housing, apparel and upkeep, transportation, medical care, entertainment, and other commodities and services.

The geographic areas for which the CPI is published and for which variances are computed are the following:

(a) The *index area* is the basic geographic area for which a fixed sample of commodities and services is priced monthly or bimonthly.

b) The four Census *regions* are Northeast, North Central, South, and West.

c) The urban *U.S.*

The CPI is a modified Laspeyres index which is a ratio of the costs of purchasing a set of items of constant quality and quantity in two different time periods. Let $IX(i,m,t,0)$ denote the index for time t for item stratum i in index area m, where 0 represents the base or reference period. We note that June 1967 is the base period for the major groups

for which variances are reported here. For Homeowner's Equivalent Rent, the base period is December 1982. Then

$$IX(i,m,t,0) = 100 * CW(i,m,t)/CW(i,m,0),$$

where $CW(i,m,t)$ and $CW(i,m,0)$ denote estimates of expenditures, termed *cost weights*, for time t and for the base or reference period 0, respectively, on stratum i.

$IX(i,m,t,0)$ may also be expressed in terms of the index for a previous period:

$$IX(i,m,t,0) = 100 * [R(i,m,t,t-1) * CW(i,m,t-1)]/CW(i,m,0),$$

where $CW(i,m,t-1)$ denotes the cost weight for stratum i at time t-1 and $R(i,m,t,t-1)$ denotes the estimate of price change, termed a one-period relative, from time t-1 to time t for stratum i, computed by:

$$R(i,m,t,t-1)=\sum_{qj} W_{iqj}(P_{iqjt}/P_{iqja})/\sum_{qj} W_{iqj}(P_{iqjt-1}/P_{iqja}),$$

where P_{iqjt} is the price of the qth quote, i.e., a sample item in a sample outlet, in the jth outlet in time period t for item stratum i; P_{iqja} is the price of the qth quote in the jth outlet in the outlet expenditure frame development reference period a; and W_{iqj} is a composite sampling weight for the qth quote in the jth outlet for item stratum i. The sum given here is over all quotes and outlets for the item stratum in the index area.

Cost weights for item strata are updated on a monthly or bimonthly basis using one period price relatives defined above. They are summed to estimate cost weights for higher level item aggregates (HLIAs) such as expenditure classes, major groups, and all items, denoted I, and higher level geographic aggregates (HLGAs), such as regions and all U.S., denoted M, by:

$$CW(I,M,t) = \sum_{m\in M} \sum_{i\in I} CW(i,m,t).$$

Thus the formula for the index for any higher level item-area aggregate is:

$$IX(I,M,t,0) = 100 * CW(I,M,t) / CW(I,M,0).$$

The CPI sample is a multi-stage probability design. The sample for any index area consists of one or more primary sampling units (PSUs) which are either metropolitan statistical areas (MSAs), or the urban parts of non-MSA counties. [See Dippo and Jacobs (1983).]. For purposes of variance computation and operational manageability, samples for all index areas are split into two or more disjoint subsets or replicates, historically called half-samples.

3. Estimating Variances of the Index and Price Change

Conditional variances of the CPI were computed using a random group estimation method. Details of this methodology are given in Leaver (1990).

The second phase of the CPI variance estimation was the estimation of the mixed covariances of the index between item strata and time periods, and estimation of the variances of the expenditure weights for December 1977. These estimates were necessary for development of the unconditional index and price change variance estimates.

Conditional mixed index covariances between item strata i and l and periods t and t-k, k=1, 2, 6, and 12, were estimated by:

$$Cov_c[IX(i,m,t,0), IX(l,m,t-k,0)] =$$

$$\frac{\sum_{j=1}^{r}\{(CW_j(i,m,t)-CW_f(i,m,t))(CW_j(l,m,t-k)-CW_f(l,m,t-k))\}}{r(r-1)CW(i,m,0)(CW(l,m,0)}.$$

Expenditure weight variances and covariances were estimated using balanced repeated replication [see Wolter (1985).] Consumer unit expenditures were collected for each index area and item stratum in either or both the Diary and Quarterly Consumer Expenditure Surveys for 1972-74. Each consumer unit was assigned a full sample weight and a set of 36 replicate weights, 18 of which were nonzero. For each replicate, region, index area and stratum, final mean expenditure estimates were computed using three steps:

(1) preliminary estimation of expenditures,
(2) composite estimation of mean expenditures,
(3) estimation of final mean expenditures, raking the mean expenditures determined in (2).

The reader is referred to *The BLS Handbook of Methods, Volume II*, (1984) for a description of these three procedures. For each item-area, the final raked mean expenditures were added across both the Diary and Quarterly Surveys. Final raked mean expenditures were then adjusted to reflect the 1977 revised item strata structure. This entailed the application of ratio adjustment factors to obtain mean expenditure estimates for newly introduced item strata. Inflation factors derived from the CPI were then applied to estimate expenditures for December 1977, denoted 7712 in the formulae below.

For each index area m, the covariance of the 7712 expenditure weights for item strata i and l was estimated using the following formula:

$$Cov[CW(i,m,7712), CW(l,m,7712)] =$$

$$[\sum_{j=1}^{36}(RME_j(i)-RME_f(i)) * (RME_j(l)-RME_f(l))]/36,$$

where $RME_j(i)$ is the estimated raked mean expenditure for a given index area m, item stratum i, and replicate j, and $RME_f(i)$ is the estimated raked full sample mean expenditure.

Second phase estimates were then combined with conditional variances to produce unconditional estimates of index and price change variances. The total variance of the index for HLIA I, index area m for time t is approximated

by the variance of a first order Taylor expansion of the ratio of cost weights at times t and 7712:

$$Var\,[IX(I,m,t,0)] = (100 / CW(I,m,0))^2 * \{Var[CW(I,m,t)] + (CW(I,m,t)/CW(I,m,7712))^2 * Var[CW(I,m,7712)] - 2 * ((CW(I,m,t)/CW(I,m,7712)) * Cov[CW(I,m,t),CW(I,m,7712)]\}.$$

Similarly, the total variance of price change from period t-k to period t,

$$PC(I,m,t,t-k) = 100 * [(IX(I,m,t,0)/IX(I,m,t-k,0))-1],$$

is approximated via the variance of a first order Taylor expansion of the ratio of cost weights at times t and t-k:

$$Var(PC(I,m,t,t-k)) = (100/CW(I,m,t-k))^2 * [Var(CW(I,m,t)) + (CW(I,m,t)/CW(I,m,t-k))^2 Var[CW(I,m,t-k)] - 2 * (CW(I,m,t)/CW(I,m,t-k)) Cov(CW(I,m,t),CW(I,m,t-k))].$$

Here $Var[CW(I,m,7712)]$ is as defined above and $Var[CW(I,m,t)]$ and $Cov[CW(I,m,t),CW(I,m,t-k)]$ are the unconditional cost weight variances and covariances for HLIA I, index area m, and times t and t-k, estimated via the equality:

$$Var\,[CW(I,m,t)] = Var[E(CW(I,m,t) / CW(I,m,7712))] + E[Var(CW(I,m,t)/ CW(I,m,7712))],$$

which gives

$$Cov[CW(I,m,t),CW(I,m,t-k)] = (CW(I,m,0)/100)^2 Cov_c [IX(I,m,t,0),IX(I,m,t-k,0)] + \sum_{i \in I} \sum_{j \in I} \{ ((IX(i,m,7712,0)IX(j,m,7712,0))^{-1} * [IX(i,m,t,0)IX(j,m,t-k,0) - Cov_c [IX(i,m,t,0),IX(j,m,t-k,0)) * Cov[CW(i,m,7712),CW(j,m,7712)]\}.$$

Index and price change total variance estimates for HLGAs were computed, assuming independence of indexes between areas, by summing cost weight covariances over index areas within the HLGA:

$$Var\,[IX(I,M,t,0)] = (100/CW(I,M,0))^2 * [Var(CW(I,M,t)) + (CW(I,M,t)/CW(I,M,7712))^2 * Var(CW(I,M,7712)) - 2 * ((CW(I,M,t)/CW(I,M,7712)) * Cov(CW(I,M,t),CW(I,M,7712))],$$ and

$$Var[PC(I,M,t,t-k)] = (100/CW(I,M,t-k))^2 [Var(CW(I,M,t)) + (CW(I,M,t)/CW(I,M,t-k))^2 Var(CW(I,M,t-k)) - 2 (CW(I,M,t)/CW(I,M,t-k))Cov (CW(I,M,t), CW(I,M,t-k))].$$

4. Findings

Figures 1-3 display the conditional and unconditional standard error of the CPI at the U.S. level for all items, seven major groups, and three housing subgroups, plotted as a function of time for the 108 months, January 1978 - December 1986. Figures 4-7 display the conditional and unconditional standard error of 12-month price change for All Items and major groups as a function of time for the same period.

It is important to note that the unconditional index standard errors for all the items shown are indistinguishable from their conditional estimates. In all cases they exhibit a gradual climb over the time period of this study. This is also reflected in the plots of index standard error versus value of the index shown for All Items, Food, and Apparel in Figure 3 in Leaver (1990). Some major groups, most notably Apparel, Entertainment, and Housing, and all three subgroups within Housing, exhibit apparent seasonal behavior in their standard errors also.

The steady increase in index variance observed in all major groups is consistent with the findings of Valliant and Miller (1989) who considered a Laspeyres-type index computed by chaining month-to-month estimates of price change. The relative variance of this estimator was shown to grow over time.

Abrupt jumps in index variance occur in Medical Care in September 1980, September 1985 and January 1986 and in Other Commodities and Services in May-June 1981 and September 1985, and Fuel and Utilities in December 1980. The nature of these jumps was investigated and attributed to real, substantial price change in sample observations in specific strata in one or two index areas in hospital and other medical care services, health insurance, personal services, tuition and other school fees, and other utilities and public services EC's. The effect of the price change in the hospital and other medical care services EC is magnified in variance calculations by the fact that the price movement for the other hospital and medical care services stratum is used, in conjunction with another medical care stratum, as a regular source of imputation for four health insurance strata. The effects of these apparent discontinuities in what otherwise is a generally smooth function can be seen in the corresponding locations in graphs of price change standard error for these major groups (see Figures 6 and 7 for 12-month price change.)

While index variance is seen to climb steadily with respect to time and value of the index, behavior of price change variance is for All Items and several major groups generally stable and uniform over the 108-month period, especially when compared to the actual fluctuation of price change estimated for that period. Figures 8-9 display price change and conditional and unconditional price change standard error versus time for All Items for 1- and 12-month lags, respectively. This stability is attributable to the high correlation of the index between months, that is, between the numerator and denominator indexes of price change estimates. Figure 10 displays the 1- and 12-month lagged correlations of the conditional All Items index. Similar results obtain for most major groups.

Price change variances for Apparel and Upkeep are a notable exception to the above in that they appear to exhibit a more seasonal variation and climb with greater speed and variability over the interval of the study. The reasons for this behavior are not entirely well understood, though it is well recognized that price collection for items in this major

group poses greater difficulties than in other major groups, especially with respect to availability of items. Other sources of nonsampling error which may be confounded in sampling error estimates include variability introduced in item substitution and imputation of reference period prices. Comparisons of numbers of usable quotes, the actual number of price change observations used in computing the price relatives for strata, as shown in Figures 11-12, however, indicate the effect of sample size in major group price change standard error. Usable quotes for Apparel and Upkeep and Entertainment are approximately one tenth the number for Food in any month.

As expected, average price change standard errors increase with lag, though not in a manner linearly proportional to the length of the interval. The average standard error for 12-month price change is approximately 2.4 times the standard error for 1-month price change and approximately 1.7 times the standard error for 2-month price change for the All Items index. Similar behavior holds for most other major groups at the U. S. All Cities and regional levels. Medical Care and Other Commodities and Services, for reasons discussed above, are notable exceptions.

As is the case with total Index variability, price change sampling variance dominates expenditure weight sampling variance in total price change variability, though to a lesser degree. The contribution of 7712 expenditure weight variance to total price change variance estimates varies by major group, ranging for 12-month price change from 6% for Apparel and Upkeep and Household Furnishings and Operations to 20% for Shelter. At the major group level, this contribution is proportionally greater for items with lower total sampling variability. Average price change, average unconditional price change standard errors, ratios of average price change standard errors to that for 1-month price change and ratios of average unconditional standard errors to conditional standard errors for 2-, 6- and 12-month lags for the U.S. All Cities index for all items and major groups and for all items for the four regions are given in Table 1.

Table 1 - Average Price Change, Average Price Change Standard Errors by Major Group for 1-, 2-, 6- and 12 Month Lags

Major Group	1-Mo Avg PC	1-Mo Avg SE	Ratio Total/ Condl	2-Mo Avg PC	2-Mo Avg SE	Ratio 2/1	Ratio Total/ Condl	6-Mo Avg PC	6-Mo Avg SE	Ratio 6/1	Ratio Total/ Condl	12-Mo Avg PC	12-Mo Avg SE	Ratio 12/1	Ratio Total/ Condl
U. S. All Cities															
All Items	0.516	0.053	1.00	1.040	0.071	1.34	1.01	3.183	0.092	1.74	1.03	6.560	0.124	2.35	1.08
Food	0.469	0.080	1.03	0.935	0.101	1.26	1.03	2.762	0.148	1.85	1.06	5.445	0.197	2.47	1.10
Housing	0.547	0.115	1.00	1.102	0.152	1.32	1.01	3.407	0.177	1.54	1.03	7.049	0.234	2.04	1.09
Shelter	0.612	0.174	1.00	1.234	0.231	1.33	1.01	3.785	0.257	1.48	1.04	7.780	0.350	2.01	1.12
Fuels & Utilities	0.545	0.179	1.01	1.106	0.247	1.38	1.01	3.577	0.390	2.18	1.03	7.689	0.512	2.86	1.08
HH Furn & Opn	0.364	0.166	0.99	0.730	0.213	1.29	1.00	2.215	0.277	1.67	1.01	4.488	0.351	2.12	1.03
Apparel & Upkeep	0.278	0.271	1.01	0.592	0.372	1.37	1.01	1.764	0.564	2.08	1.01	3.474	0.746	2.75	1.03
Transportation	0.500	0.061	1.00	1.014	0.084	1.37	1.00	3.158	0.129	2.12	1.02	6.780	0.177	2.90	1.07
Medical Care	0.713	0.117	1.01	1.433	0.170	1.46	1.01	4.384	0.309	2.65	1.04	9.042	0.463	3.97	1.09
Entertainment	0.449	0.183	1.01	0.905	0.252	1.37	1.01	2.748	0.398	2.17	1.04	5.674	0.558	3.04	1.08
Other C & S	0.643	0.103	1.01	1.296	0.146	1.42	1.01	3.998	0.254	2.47	1.03	8.244	0.378	3.68	1.07
Region 1 - Northeast															
All Items	0.525	0.105	1.00	1.057	0.140	1.33	1.01	3.221	0.183	1.74	1.04	6.669	0.254	2.42	1.10
Region 2 - North Central															
All Items	0.480	0.086	1.00	0.967	0.116	1.34	1.01	2.960	0.153	1.77	1.02	6.070	0.203	2.35	1.06
Region 3 - South															
All Items	0.512	0.098	0.99	1.033	0.131	1.34	1.01	3.158	0.171	1.75	1.04	6.496	0.231	2.36	1.09
Region 4 - West															
All Items	0.555	0.128	0.99	1.121	0.177	1.38	1.01	3.449	0.218	1.70	1.02	7.114	0.291	2.27	1.06

It is also useful to note that standard errors of 1-month price change for All Items are, in times of low inflation, on the same order as the price change estimates themselves. It is only for 6- and 12-month lags that in most cases the standard errors of small price changes do not overwhelm the price change estimates themselves.

5. Conclusions

The final estimates of standard error for the CPI and price change are little different from their conditional values. That is, price change sampling variance dominates expenditure estimate sampling variance in total price change variability. As was noted earlier in Leaver (1990),

these estimates raise important issues regarding the sample design and index estimation for the survey. Extreme changes in price in a few quotes can have a dramatic effect on the variance of a major group and lower level index or price change estimators. This is particularly the case when significant imputation occurs. This points to the need to scrutinize imputation assumptions and account for effective stratum relative importance in sample allocation.

Standard error of price change appears to be a fairly stable measure over the period of this study, which included periods of extremely high and extremely low inflation. This is a comforting finding, which is largely attributable to the high correlation of lagged indexes over time. However, the similarities in magnitude of standard errors and price change for small values of price change, particularly for 1- and 2- month lags, indicate the need for caution in inferring level and direction of price change trends.

6. Acknowledgments

The authors would like to acknowledge Michael P. Cohen for his development of the computing formulae for total index and price change variances, and Edward Pratt, David L. Swanson, Rita Smith, Donald Jones and John Dash for their assistance in production processing of variance estimates. Daniel H. Ginsburg provided valuable information in researching particular instances of price change. Special thanks go to Gary R. Zion for his help in the graphical display of these results. The authors also wish to acknowledge the late Curtis A. Jacobs for his support and encouragement in this effort, and to thank Richard Valliant, Brian Hedges, Janet Williams, and James Branscome for their careful reading of this and earlier drafts of this paper and for their helpful comments.

7. References

Bureau of Labor Statistics, *BLS Handbook of Methods, Volume II*, U.S. Government Printing Office, 1984, 20-21.

Bureau of Labor Statistics, *BLS Handbook of Methods*, U.S. Government Printing Office, 1988, 154-208.

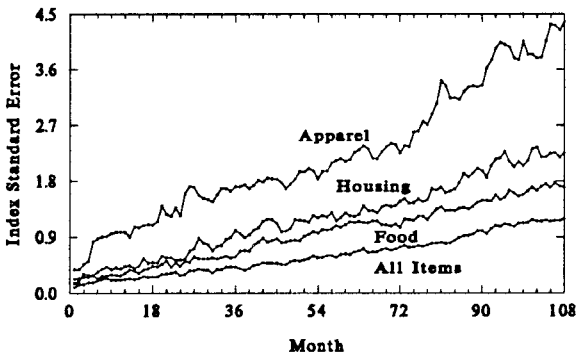
Dippo, Cathryn S., and Jacobs, Curtis A., "Area Sample Redesign for the Consumer Price Index," *Proceedings of the Survey Research Methods Section*, American Statistical Association, 1983, 118-23.

Leaver, Sylvia G., "Estimating Variances for the U.S. Consumer Price Index for 1978-1986," *Proceedings of the Survey Research Methods Section*, American Statistical Association, 1990, 290-295.

Valliant, R. and Miller, S. M., "A Class of Multiplicative Estimators of Laspeyres Price Indexes," *Journal of Business and Economic Statistics*, 77, 1989, 387-394.

Wolter, K., *Introduction to Variance Estimation*, Springer-Verlag, New York, 1985.

Figures 1-2. Index Standard Error vs Time for All Items, 7 Major Groups



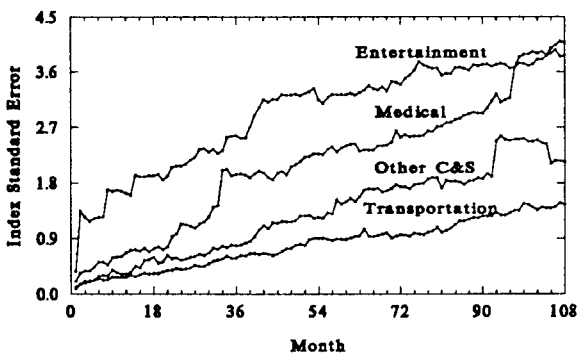


Figure 3. Index Standard Error vs Time for 3 Housing Subgroups

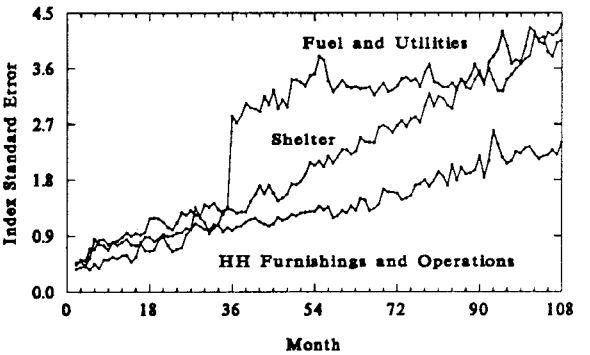


Figure 4. 12-Month Price Change Standard Error vs Time for All Items and Housing

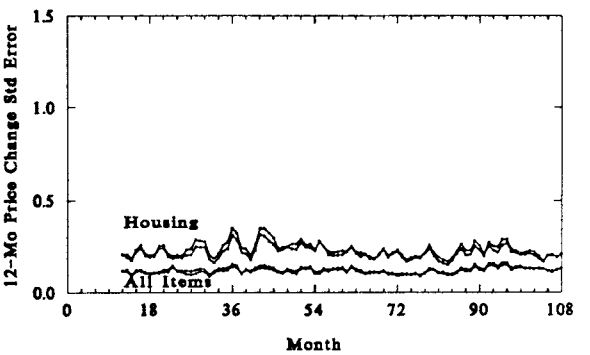


Figure 5. 12-Month Price Change Standard Error vs Time for Food and Apparel

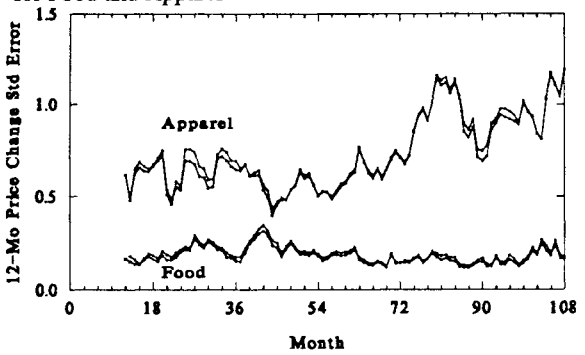


Figure 6. 12-Month Price Change Standard Error vs Time for Transportation, Medical Care

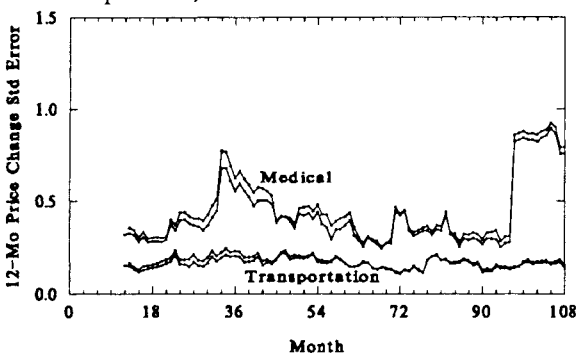


Figure 7. 12-Month Price Change Standard Error vs Time for Entertainment, Other C&S

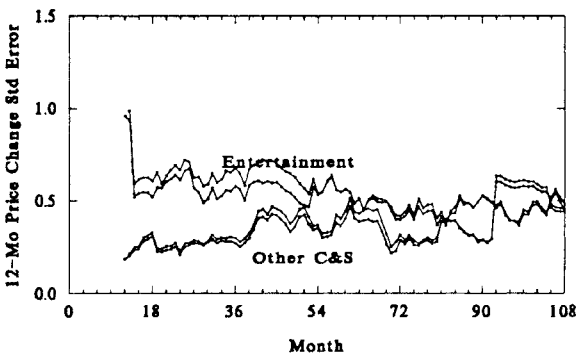


Figure 8. 1-Month Price Change, Price Change Standard Error vs Time for All Items

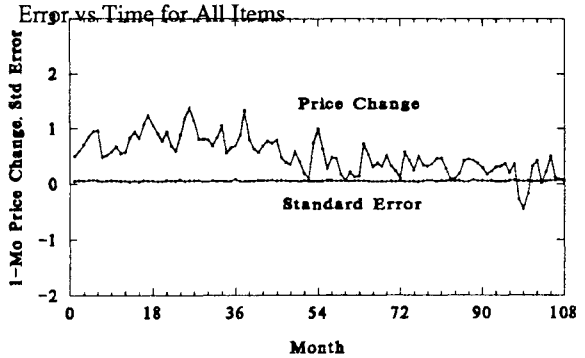


Figure 9. 12-Month Price Change, Price Change Standard Error vs Time for All Items

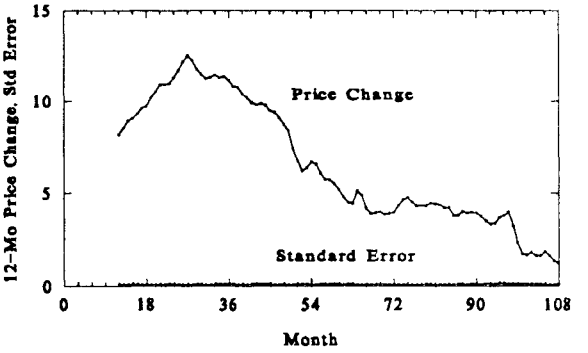
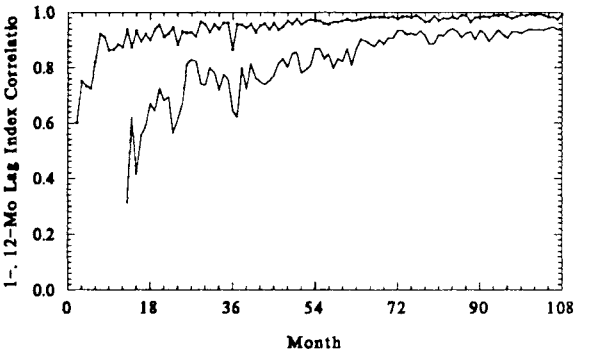


Figure 10. Index Correlation vs Time for All Items, for 1- and 12-Month Lags



Figures 11-12. Usable Quotes vs Time for Selected Major Groups