SUPPORTING STATEMENT
Consumer Price Index Commodities and Services

B.   DESCRIPTION OF INFORMATION COLLECTIONS EMPLOYING
       STATISTICAL METHODS

1.   Universe and Sample Size Summary

Because of the complexity, importance and diversity of its universe, the construction of the CPI requires a complex set of statistical techniques and samples. Conceptually, the potential respondent universe of price quotations for the CPI is the total set of prices, placed in one-to-one correspondence to the total set of purchases of all urban consumers. The sample for ongoing pricing for the C&S portion of the CPI (including outlets selected but not directly contacted, e.g., using the SABRE system to price airfares) is approximately 39,073 outlets with 135,551 price quotations.

The outlet response rate for ongoing pricing is 96.0 percent per month over time period from November 2006 to October 2007. The four percent non response in outlets is due to refusals (we remove outlet from the sample) or outlets being temporarily unavailable for pricing. The overall response rate at initiation is 90 percent (data obtained or pending); that is, 10 percent of the outlets sent out are terminated during initiation.   The following table presents response rates for outlets under going CPI initiation during the two most recent initation cycles, October 2006 and February 2007:

Type of Response	Percent
Data obtained	79.6
Data pending	10.5
Refusal	1.7
No CPI items available	3.6
Out-of-business, out of scope, outlet moved	2.8
Unable to locate	1.8

2.   Collection Procedures

2.i. Description of Sampling Methodology

A multi-stage sampling process is employed for the CPI. The four main stages of selection are: (1) the sampling of geographic areas, (2) the sampling of outlets within the geographic areas, (3) the sampling of item groups (ELI's) to be priced in the outlets, and (4) the sampling of items from each ELI in each outlet.

(a) BLS selects Primary Sampling Units (PSU's) or geographic areas for pricing. The sample pricing areas were derived from a stratified design using a controlled selection procedure that provided for the selection of one sample area from each stratum with a control on the distribution of PSU's by state. The initial stratification for the PSU design was based on such variables as percent black consumer units, percent fuel oil heated housing units, percent electric heated housing units, mean wage and salary income per consumer unit, and mean interest and dividend income per consumer unit.

(b) Each year BLS systematically re-selects a portion of the sample of outlets and quotes such that over a four-year period most C&S sample outlets have a chance to be replaced. Not only does this re-establish the distribution of the sample, incorporate new outlet construction and

reflect shifts in outlet preferences, but it also allows many respondents to rotate out of the sample. Thus, all respondents are not indefinitely retained in the sample.

The outlet sampling frames are constructed from several sources. The primary source for all food and the majority of the other C&S items is the Telephone Point of Purchase Survey (TPOPS). The TPOPS provides coverage for 60.3 percent of all consumption expenditures for the CPI-U, as of December 2006. Renter and owner-occupied housing account for 29.8 percent. The remaining 9.9 percent of consumption expenditures are covered from a variety of sampling frames constructed by BLS or obtained from other sources.

The TPOPS is a computer assisted telephone collection effort, used to identify a universe of outlets from which CPI sample outlets are selected, and is conducted by the Census Bureau for BLS. TPOPS is made up of 214 purchase categories of goods and services, e.g., prescription drugs. Under TPOPS, during each quarter of the year, in rotating groups of PSU/purchase category groups, households are asked to identify the amount of their expenditures and the names and addresses of the outlets where purchases were made. Samples of outlets for pricing are selected from the TPOPS generated frames using a systematic sampling procedure with each outlet having a probability of selection proportional to the expenditures reported for it on the TPOPS.

(c) The sampling frames from which the item sample market baskets are derived are constructed using data from the most current two years of the Continuing Consumer Expenditure Survey, which is an ongoing survey. Each year as we rotate a portion of the outlet sample we also resample the ELIs. With data from these surveys assembled into the CPI item classification structure, we select the sample of ELI's using a stratified random selection procedure with each ELI having a probability of selection proportional to the expenditures reported for it on the Continuing Consumer Expenditure Surveys.

(d) The BLS Washington Office merges the sample of ELIs with the appropriate sample of outlets. BLS field representatives then initiate the new outlets and select the specific unique items to be priced within each ELI by following an outlet based multistage probability proportional to sales methodology.

2.ii. Description of Estimation Methodology History
A price index constructed using geometric means more closely approximates a true cost-of-living index than does the Laspeyres. This occurs because the geometric means formula, unlike the Laspeyres formula, implicitly assumes that product substitution takes place when relative prices change. The geometric means formula assumes that relative expenditures are kept constant over time.

The Laspeyres index formula in concept simply measures the change in the weighted arithmetic mean of prices. As a fixed-weight index, the Laspeyres formula assumes that consumers do not change the amount of each item purchased as relative prices change. In reality, as relative prices change, consumers do switch toward items whose prices become relatively lower.

Based on December 2006 CPI-U relative importances, 60 percent of the CPI is calculated using a Geometric mean formula and 40 percent is based on the Laspeyres index formula. (The Laspeyres portion is composed of Rent, 6 percent, Owners' equivalent rent, 24 percent, and C&S items, 10 percent) Also note that C&S items accounts for 70 percent of the CPI-U weight.

All C&S stratum indexes are calculated using a geometric formula, except for those listed below. Demand elasticity studies led BLS to conclude that the Laspeyres index formula would yield the least biased measure of price change for these items.

C&S Components retaining the Laspeyres (arithmetic mean) Formula		
Housing at school, excluding board;	Electricity; Utility piped gas service;	Residential water and sewerage maintenance;
Telephone services, local charges;	State and local registration, license and motor vehicle property tax;	Physicians' services;
Dental services; Services by other medical professionals;	Hospital services;	Nursing homes and adult daycare.

Price relatives.

The price relative for each basic item-area for commodities and services (C&S) using the **geometric mean**-is based on the formula:

$$_{a,i}R^G_{[t;t-1]} = \prod_{j \varepsilon a,i} \left(\frac{P_{j,t}}{P_{j,t-1}} \right)^{\left(W_{j,POPS} / \sum_{k \varepsilon a,i} W_{k,POPS} \right)}$$

The price relative for each basic item-area for commodities and services (C&S) using **Laspeyres**-is based on the formula:

$$_{a,i}R^L_{[t;t-1]} = \frac{\sum_{j \varepsilon a,i} \left(W_{j,POPS} / P_{j,POPS} \right) P_{j,t}}{\sum_{j \varepsilon a,i} \left(W_{j,POPS} / P_{j,POPS} \right) P_{j,t-1}}$$

$_{a,i}R^G_{[t;t-1]}$ and $_{a,i}R^L_{[t;t-1]}$, are, respectively, the geometric and Laspeyres price relatives for area-item combination, a,i, from the previous period, t-1 (either 1 month or 2 months ago), to the current month, t;

$P_{j,t}$ is the price of the jth observed item in month t for area-item combination a,i;

$P_{j,t-1}$ is the price of the same item in time t-1;

$P_{j,POPS}$ is an estimate of the item j's price in the sampling period when its POPS was conducted; and

$W_{j,POPS}$ is item j's weight in the POPS,

The product in the geomenas formula and sums in the Lasperyres formual are taken over all useable quotes in area-item combination *a, i*.[1]

The price relative for each basic item-area for commodities and services (C&S) using **Laspeyres**-is based on the formula:

---

[1] For additional details regarding the C-CPI-U, please see BLS Handbook of Methods, Chapter 17, The Consumer Price Index, pages 21-22. http://stats.bls.gov/opub/hom/homch17_a.htm

$$\frac{\sum (FEPX \& CP_t / FSBP\_WTX) * FNLW \& CP}{\sum (FEPX \& CP_{t-n} / FSBP\_WTX) * FNLW \& CP}$$

for usable quotes where:

$FEPX\&CP_t$ is the full sample effective price with tax in
   collection period t;
$FEPX\&CP_{t-n}$ is the previous on-cycle full sample effective
   price with tax (t-1 for monthly quotes, t-2 for
   bimonthly quotes);
FSBP_WTX is the full sample base period price with tax;
FNLW&CP is the final weight, and all summations are for the
   item-area.

The price relative for each basic item-area for commodities and services (C&S) using the
geometric mean-is based on the formula:

$$\left[\prod (GEPX \& CP_t / GEPX \& CP_{t-n})^{GNLW \& CP} \right]^{\left(1 / \sum GNLW \& CP \right)} \quad (1)$$

which is mathematically equivalent to

$$\exp \left[\frac{1}{\sum GNLW \& CP} * \sum (GNLW \& CP * \log(GEPX \& CP_t / GEPX \& CP_{t-n})) \right] \quad (2)$$

where log represents the natural logarithm, the products and/or summations are for the item-area
and the following geomeans variables are used: [it is not self-evident why the raw ingredients for
the Lasp. And geomean indexes should differ; some explanation and reference should be given.
Also, since these formulas as given are rather cryptic, for example the lack of indication of what
the sums are over, some reference should be given that would supply further details]
$GEPX\&CP_t$ is the geomeans full sample effective price with
   tax in collection period t;
$GEPX\&CP_{t-n}$ is the geomeans previous on-cycle full sample
   effective price with tax (t-1 for monthly quotes,
   t-2 for bimonthly quotes); and
$GNLW \& CP$ is the geomeans final weight
Index calculation. When aggregating together price relatives above the elementary index level,
the Laspeyres formula is used exclusively implying no substitution across strata in the CPI.

C-CPI -U.[2] In mid-2002, BLS began publishing a Chained Consumer Price Index for All Urban
Consumers (C-CPI-U).   The C-CPI-U is a monthly-chained index that uses a Tornqvist formula
to aggregate indexes. This index is designed to be a closer approximation to a "cost-of-living
index" than the present measures. By utilizing expenditure data in adjoining periods, it reflects
consumer substitution across item categories in response to relative prices. The use of
expenditure data for both a base period and the current period to average price change across
item categories distinguishes the C-CPI-U from the existing CPI measures. Expenditure data
required for the C-CPI-U calculations are available only with a lag. Thus, the C-CPI-U, unlike

---

[2] For additional details regarding the C-CPI-U, please see BLS Handbook of Methods, Chapter
17, The Consumer Price Index, pages 33-38. http://stats.bls.gov/opub/hom/homch17_a.htm

the CPI-U and CPI-W, is issued first in preliminary form and then subject to two subsequent revisions. No additional data collection is required to support the publication of the C-CPI-U.

2.iii. Degree of Accuracy Required

The statute mandating the CPI does not specify a required precision or accuracy for the index. BLS requires that the precision of the CPI be maximized given the total cost constraint imposed by the authorized spending level. BLS developed an allocation model to examine relative efficiencies of various alternative sample designs. The objective of the allocation process is to determine values for all sample design parameters which will minimize the variance of price change for the CPI at the U.S. level subject to the total cost constraint of the CPI budget. The model uses a variance function to project the variance of price change given a set of sample design parameters. It also has a cost function to project the annual cost given a set of values for the sample design parameters. A non-linear programming technique is used to determine the set of values for the sample design parameters which minimizes the variance of price change given a cost constraint. (For a complete description of the allocation process, see: Jacobson, Shawn, Leaver, Sylvia G. and Swanson, David C. (1998), "Choosing a Variance Computation Method for The Revised Consumer Price Index," Proceedings of the Business and Economics Statistical Section, American Statistical Association, 131-136, and Swanson, David C., (1999), "Variance Estimates for Changes in the Consumer Price Index, January 1998-December 1998," CPI Detailed Report, December 1999, Washington, D.C.: U.S. Government Printing Office, 7-20.

The estimate of the CPI-U median standard error for 12-month intervals from December 2005 through December 2006 was 0.16 for All Items.

Effective with the release in February 2007 of the January 2007 Consumer Price Index (CPI), the Bureau of Labor Statistics (BLS) began to publish its consumer price indexes rounded to three decimal places. Percent changes will be calculated from the three decimal place indexes. Those percent changes will continue to be published to one decimal place.

2.iv. Special Sampling -- Sampling of Time

The outlet samples of each PSU are divided into three pricing periods. Each outlet is designated for pricing during a specified period of the month. Therefore, a given item is priced at different times in different outlets in order to average out possible systematic differences between one time period of the month and another and to ensure that there is a full month between pricing for each monthly priced outlet or a full two months between pricings for bi-monthly collected outlets.

2.v. Use of Periodic Data Collection Cycles
Although BLS publishes monthly estimates of the CPI, prices for about 58 percent of the total covered expenditures are collected monthly in all sampling areas. Of the 58 percent priced monthly, 30 percent reflects rent and owner's equivalent rent and 28 percent C&S items.

Regarding just the C&S portion (70 percent) of the total CPI expenditure weight, 29 percent is collected monthly and 41 percent is collected bimonthly. The monthly priced C&S items include Food at home, Lodging away from home, Tenants insurance, Household fuels, Motor fuels, Motor vehicle parts, equipment and fees, Recreational reading materials, Education, Postage and delivery, Telephone services, and Tobacco products. (Note, in the three largest consolidated areas, New York, Chicago and Los Angeles all sampled items are priced monthly.) Other commodities and services are priced bi-monthly ("even" cycle--February, April, June, August, October and December or "odd" cycle--January, March, May, July, September and November.)

3. Methods of Maximizing Response

BLS utilizes several techniques to insure that adequate sample sizes are maintained for estimating the CPI. Initial sample sizes are larger than the desired sample sizes to cover initial

non-responses, i.e., out-of-business, out-of-scope, refusal, sample items not available, and unable to locate. In rare circumstances, if the sample of outlets is deemed insufficient, we continue pricing the current sample. Additionally, the CPI is exploring methods to improve its sample maintenance system that would allow us to offset sample loss in the primary sample. Benefits include improved sample sufficiency and adequacy of CPI indexes, and a reduction in variances

4. Statistical Responsibility

W. John Layng, Assistant Commissioner, Division of Consumer Prices and Price Indexes, Office of Prices and Living Conditions of BLS (telephone: 202-691-6950) is the CPI program manager and has overall responsibility for the CPI. Bill Johnson, Chief of the Survey and Analysis Research Branch of the Price Statistical Methods Division of the Office of Prices and Living Conditions (Telephone 202-691-6912) has reviewed and approved the statistical methodology for the survey design. BLS's Office of Field Operation will collect all data. The Division of Consumer Price Computer Systems of the Office of Technology and Survey Processing of BLS will process the data.

5. CPI Revision Geographic Area Update

As mentioned previously, the CPI's revision of its geographical areas is based on the assumption that the CPI will receive approval of its budgetary requests to fund such activity. If funding is not approved to support the geographic rotation, the CPI will pursue the fall back activity and continue pricing and initiating with its current 87 PSU design.

A new sample of geographic areas, which includes 31 new primary sampling units (PSUs), has been selected based on the 2000 Decennial Census. The CPI geographical revision is scheduled to begin with the initiation of data in February 2009 and the publication of the January 2010 index. The introduction of new geographical areas will occur on a continuous basis over a number of years as new areas are introduced and old areas are dropped. Once the revision is complete (scheduled for Jan 2014) the CPI will have 86 PSUs, down from the current 87. Thus the overall burden to respondents is expected to be minimal.

Applicable Performance Goal: Improve the effectiveness of information and analysis on price changes by improving the timeliness, accuracy, and relevance of the Consumer Price Index (CPI). The performance goal supports directly DOL Strategic Goal 1, A Prepared Workforce. The focus of this goal is on producing statistics and improving the quality of the economic data, not only in terms of statistical validity, but also in terms of geographic detail, industry detail, demographic detail, and subject matter.

In the past, the program has met its performance goal, including its relevancy, timeliness, accuracy, improvement, and efficiency targets. The Administration's assessment, which rated BLS as "effective" in 2005, found that the BLS fills a unique role in the development of national labor-related statistics, and that its design for data collection and reporting is sound.

Small Sample Bias
BLS research shows that a small sample of price changes used to calculate the basic components of the index will yield an upward bias in the inflation estimator of 0.1 percentage points or more per year. Reducing this bias requires an increase in sample through additional data collection and/or through pooling sample across basic components or areas. BLS has requested funding to continue to study and address this source of bias.