Consumer Price Indices on the Economics Data Program Website

Data Series Descriptions

Economics Data Program
Pacific States Marine Fisheries Commission

7600 Sand Point Way NE, Building 4
Seattle, WA 98115
efin@psmfc.org " http://www.psmfc.org/efin
(206) 526-4251 voice " (888) 421-4251 toll free
(206) 526-4074 fax

Consumer Price Indices
Data Series Descriptions

Table of Contents

This document was created to assist users of CPI data from the Economics Data Program Website. It contains the following information:

General Information

Data Retrieval Location

This data was retrieved from the Pacific States Marine Fisheries Commission's Economics Data Program at http://www.psmfc.org/efin.

Contact Information

Please direct any questions, suggestions or other comments to:

Geana Tyler
Assistant Fisheries Economics Data Manager
geana_tyler@psmfc.org
(206) 526-6683

CPI Data Sources

All CPI data are from the US Department of Labor, Bureau of Labor Statistics. They are available at http://146.142.4.24/cgi-bin/srgate using the BLS Series ID provided in the description of the data series.

All Urban Consumers

All CPI's available for download from the Economics Data Program are based on all urban consumers.

CPI's based on Urban Wage Earners and Clerical Workers may be obtained at http://146.142.4.24/cgi-bin/srgate. To obtain a CPI based on urban wage earners and clerical workers, use the BLS Series ID for the item you want replacing the leading "CU" with "CW". For instance, the BLS Series ID for a CPI for urban wage earners and clerical workers in the Western Region would be CWUR0400SA0 (instead of the CPI for all urban consumers CUUR0400SA0.) For additional help, follow the Series ID Formats link.

Seasonal Adjustment

All CPI's available for download from the Economics Data Program are NOT seasonally adjusted.

Seasonally adjusted CPI's may be obtained at http://146.142.4.24/cgi-bin/srgate. To obtain a seasonally adjusted CPI, use the BLS Series ID for the item you want replacing the second "U" with an "S". For instance, the BLS Series ID for a seasonally adjusted CPI for all urban consumers in the Western Region would be CUSR0400SA0 (instead of the unadjusted CUUR0400SA0.) For additional help, follow the Series ID Formats link.

Periodicity of Data

All CPI's on the Economics Data Program website are available annually and either monthly, bi-monthly, or semiannually. Both types of data (annual and (bi)monthly/semiannual) are stored within each downloadable file. Use the following selection procedures if you only wish to view one type of data.

(Bi)Monthly vs. Annual Data

- Select period = M13 for annual data.
- Select period <> M13 for (bi)monthly data.

Semi-Annual vs. Annual Data

- Select period = S3 for annual data.
- Select period <> S3 for semi-annual data.

All Items

CPI's are available for download based on All Items or based on Energy Items Only.

CPI's based on other subsets of items may be obtained at http://146.142.4.24/cgi-bin/srgate. To obtain a CPI based on another subset of items, use the BLS Series ID for the item you want replacing "A0" (or "A0E" for energy items only) at the end of the series with the item code you want. For additional help, follow the Series ID Formats link.

Energy Items Only

CPI's are available for download based on All Items or based on Energy Items Only.

CPI's based on other subsets of items may be obtained at http://146.142.4.24/cgi-bin/srgate. To obtain a CPI based on another subset of items, use the BLS Series ID for the item you want replacing "A0E" (or "A0" for all items) at the end of the series with the item code you want. For additional help, follow the Series ID Formats link.

Data Series Descriptions for Consumer Price Indices: All Urban Consumers: All Items

by region

Anchorage AK

BLS Series ID CUUSA427SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available semi-annually (S)
For the metropolitan statistical area (MSA) of Anchorage, AK (A427)
Base period: 1982-84=100 (S)
All items (A0)

Los Angeles-Riverside-Orange County, CA

BLS Series ID CUURA421SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available monthly (R)
For the metropolitan statistical area (MSA) of Los Angeles-Riverside-Orange County, CA (A421)
Base period: 1982-84=100 (S)
All items (A0)

Portland-Salem, OR-WA

BLS Series ID CUUSA425SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available semi-annually (S)
For the metropolitan statistical area (MSA) of Portland-Salem, OR-WA (A425)
Base period: 1982-84=100 (S)
All items (A0)

San Diego, CA

BLS Series ID CUUSA424SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available semi-annually (S)
For the metropolitan statistical area (MSA) of San Diego, CA(A424)
Base period: 1982-84=100 (S)
All items (A0)

Seattle-Tacoma-Bremerton, WA

BLS Series ID CUUSA424SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available semi-annually (S)
For the metropolitan statistical area (MSA) of Seattle-Tacoma-Bremerton, WA (A423)
Base period: 1982-84=100 (S)
All items (A0)

San Francisco-Oakland-San Jose, CA

BLS Series ID CUURA422SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available bi-monthly (R)
For the metropolitan statistical area (MSA) of San Francisco-Oakland-San Jose, CA (A422)
Base period: 1982-84=100 (S)
All items (A0)

Western Region

BLS Series ID CUUR0400SA0

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available monthly (R)
For the BLS Western Region, including CA, OR, WA, AK, AZ, HI, ID, NV (0400)
Base period: 1982-84=100 (S)
All items (A0)

Data Series Descriptions for Consumer Price Indices: All Urban Consumers: Energy Items Only

by region

Los Angeles-Riverside-Orange County, CA

BLS Series ID CUURA421SA0E

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available monthly (R)
For the metropolitan statistical area (MSA) of Los Angeles-Riverside-Orange County, CA (A421)
Base period: 1982-84=100 (S)
Energy Items Only (A0E)

Seattle-Tacoma-Bremerton, WA

BLS Series ID CUUSA424SA0E

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available semi-annually (S)
For the metropolitan statistical area (MSA) of Seattle-Tacoma-Bremerton, WA (A423)
Base period: 1982-84=100 (S)
Energy Items Only (A0E)

San Francisco-Oakland-San Jose, CA

BLS Series ID CUURA422SA0E

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available monthly (R)
For the metropolitan statistical area (MSA) of San Francisco-Oakland-San Jose, CA (A422)
Base period: 1982-84=100 (S)
Energy Items Only (A0E)

Western Region

BLS Series ID CUUR0400SA0E

CPI based on all urban consumers (CU)
Not seasonally adjusted (U)
Available monthly (R)
For the BLS Western Region, including CA, OR, WA, AK, AZ, HI, ID, NV (0400)
Base period: 1982-84=100 (S)
Energy Items Only (A0E)

Bureau of Labor Statistics Information

IMPORTANT NOTE

The BLS Overview of the CPI and the BLS Handbook of Methods have been added to this file for your convenience. These documents were downloaded directly from the BLS website at the URL's below on 16 September 1999.

You are encouraged to view the most recent version of these documents on the BLS website before including either document in your citations. The URL's are listed below.

If you do discover that either of these documents is not the most recent version, we would greatly appreciate if you would notify the Economics Data Program. The Economics Data Program can be reached via email at efin@psmfc.org, via phone at (206) 526-4251, (206) 526-6683 or toll free at (888) 421-4251.

BLS Resources on the Web

The BLS Overview contained in this section was obtained at http://www.bls.gov/cpiovrvw.htm.

The BLS Handbook of Methods, Chapter 17: The Consumer Price Index was obtained at http://stats.bls.gov/opub/hom/homch17_a.htm.

Additionally, you may find the following websites useful when using the CPI:

- The BLS CPI Homepage is located at http://www.bls.gov/cpihome.htm
- The BLS Answers to Frequently Asked Questions is located at http://www.bls.gov/cpifaq.htm
- BLS Contact Information is located at http://www.bls.gov/cpicont.htm
- The BLS Homepage is located at http://stats.bls.gov/blshome.htm
- The BLS form-based querying system is located at http://stats.bls.gov/sahome.html

[Accessibility Information]

Consumer Price Indexes

Overview

Consumer Price Indexes

Contents

Data available

- Price indexes are available for the U.S., the four Census regions, size of city, cross-classifications of regions and size-classes, and for 26 local areas. Indexes are available for major groups of consumer expenditures (food and beverages, housing, apparel, transportation, medical care, recreation,education and communications, and other goods and services), for items within each group, and for special categories, such as services.
- Monthly indexes are available for the U.S., the four Census regions, and some local areas. More detailed item indexes are available for the U.S. than for regions and local areas.
- Indexes are available for two population groups: a CPI for All Urban Consumers (CPI-U) which covers approximately 87 percent of the total population and a CPI for Urban Wage Earners and Clerical Workers (CPI-W) which covers 32 percent of the population.
- Some series, such as the U.S. City Average All items index, begin as early as 1913.

Coverage

- The CPI represents changes in prices of all goods and services purchased for consumption by urban households. User fees (such as water and sewer service) and sales and excise taxes paid by the consumer are also included. Income taxes and investment items (like stocks, bonds, and life insurance) are not included.
- The CPI-U includes expenditures by urban wage earners and clerical workers, professional, managerial, and technical workers, the self-employed, short-term workers, the unemployed,

retirees and others not in the labor force. The CPI-W includes only expenditures by those in hourly wage earning or clerical jobs.

Sources of data

- Prices for the goods and services used to calculate the CPI are collected in 87 urban areas throughout the country and from about 23,000 retail and service establishments. Data on rents are collected from about 50,000 landlords or tenants.
- The weight for an item is derived from reported expenditures on that item as estimated by the Consumer Expenditure Survey.

Reference period

- Prices are taken throughout the month.

Forms of publication

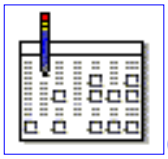
- Monthly news release. Consumer Price Index. Electronic access available.
- Historical data in Handbook of Labor Statistics. Electronic access available.
- Diskettes
- LABSTAT database.

Uses

- As an economic indicator. As the most widely used measure of inflation, the CPI is an indicator of the effectiveness of government policy. In addition, business executives, labor leaders and other private citizens use the index as a guide in making economic decisions.
- As a deflator of other economic series. The CPI and its components are used to adjust other economic series for price change and to translate these series into inflation-free dollars.
- As a means for adjusting income payments. Over 2 million workers are covered by collective bargaining agreements which tie wages to the CPI. The index affects the income of almost 80 million people as a result of statutory action: 47.8 million Social Security beneficiaries, about 4.1 million military and Federal Civil Service retirees and survivors, and about 22.4 million food stamp recipients. Changes in the CPI also affect the cost of lunches for the 26.7 million children who eat lunch at school. Some private firms and individuals use the CPI to keep rents, royalties, alimony payments and child support payments in line with changing prices. Since 1985, the CPI has been used to adjust the Federal income tax structure to prevent inflation-induced increases in taxes.

Major research in progress

- Continuing research on technical improvements in the calculation of the CPI
- Continuing work on the next major weight revision of the CPI



CPI Home Page

BLS Home Page

Sharon Gibson
Bureau of Labor Statistics
gibson_s@bls.gov
Last modified: July 07, 1998
URL: http://stats.bls.gov/cpiovrvw.htm

Chapter 17.
The Consumer Price Index

Part I. The Index in Brief

The Consumer Price Index (CPI) measures the average change in the prices paid by urban consumers for a fixed market basket of goods and services. The Bureau of Labor Statistics (BLS) calculates the CPI monthly and publishes it about 2 weeks after the end of the month to which it refers.

BLS calculates the CPI for two population groups, one consisting only of wage earners and clerical workers and the other consisting of all urban consumers.[1] The Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W) is a continuation of the historical index that was introduced well over a half-century ago for use in wage negotiations. As new uses were developed for the CPI in recent years, the need for a broader and more representative index became apparent. The Consumer Price Index for All Urban Consumers (CPI-U) introduced in 1978 is representative of the buying habits of about 80 percent of the non-institutional population of the United States, compared with 32 percent represented in the CPI-W. The methodology for producing the index is the same for both populations and is described in detail in part II of this chapter.

Background

The Consumer Price Index was initiated during World War I, when rapid increases in prices, particularly in shipbuilding centers, made such an index essential for calculating cost-of-living adjustments in wages. To provide appropriate weighting patterns for the index, so that it would reflect the relative importance of goods and services purchased by consumers, studies of family expenditures were conducted in 92 industrial centers in 1917-19. Periodic collection of prices was started, and, in 1919, the Bureau of Labor Statistics began publication of separate indexes for 32 cities. Regular publication of a national index, the U.S. city average, began in 1921, and indexes were estimated back to 1913.[2]

IN THIS CHAPTER

[1] The all-urban-consumer population consists of all-urban households in Metropolitan Statistical Areas (MSA's) and in urban places of 2,500 inhabitants or more. Nonfarm consumers living in rural areas within MSA's are included, but the index excludes rural consumers and the military and institutional population. The urban wage earner and clerical worker population consists of consumer units with clerical workers, sales workers, craft workers, operative, service workers, or laborers. More than one-half of the consumer unit's income has to be earned from the above occupations, and at least one of the members must be employed for 37 weeks or more in an eligible occupation.

[2] Collection of food prices back to 1890 had been initiated in 1903. During the course of the 1917-19 expenditure survey, retail prices for other articles were collected in 19 cities for December of each year back to 1914 and in 13 other cities back to December 1917 only, Retail prices of food and wholesale prices of other items were used to estimate price change from 1914 back to 1913.

Because people's buying habits had changed substantially, a new study was made covering expenditures in the years 1934-36, which provided the basis for a comprehensively revised index introduced in 1940.

During World War II, when many commodities were scarce and goods were rationed, the index weights were adjusted temporarily to reflect these shortages. In 1951, the BLS again made interim adjustments, based on surveys of consumer expenditures in seven cities between 1947 and 1949, to reflect the most important effects of immediate postwar changes in buying patterns.[3]

The first comprehensive postwar revision of the index was completed in January 1953, using weights from the 1950 expenditure survey.[4] At that time, not only were the weighting factors, list of items, and sources of price data updated (appendix 1), but many improvements in pricing and calculation methods were introduced. Medium-size and small cities were added to the city sample to make the index representative of prices paid by all urban wage earner and clerical-worker families.

Another revision, completed in 1964, introduced new expenditure weights based on spending patterns in 1960-61 of single persons as well as families, and updated samples of cities, goods and services, and retail stores and service establishments.[5]

The 1978 revision reflected spending patterns based upon the surveys of consumer expenditures conducted in 1972-74. A new and expanded 85-area sample was selected based on the 1970 Census of Population. The Point-of Purchase Survey (POPS) was also introduced. POPS eliminated reliance on outdated secondary sources for screening samples of establishments or outlets where prices are collected. A new store-specific approach to the item selection process was also introduced, as was a second index, the more broadly based CPI for All Urban Consumers. The CPI-U took into account the buying patterns of professional and salaried workers, part-time workers, the self-employed, the unemployed, and retired people, in addition to wage earners and clerical workers.[6]

In January 1983, the Bureau changed the way in which homeownership costs were measured.[7] A rental equivalence method replaced the asset-price approach to home-ownership costs for the CPI-U. In January 1985, the same change was made in the CPI-W. The central purpose of the change was to separate shelter costs from the investment component of homeownership so that the index would reflect only the cost of shelter services provided by owner-occupied homes.

The most recent revision of the CPI, completed in 1987, further improved sampling, data collection, processing, and statistical estimation. This revision stressed techniques that would make the production and calculation of the CPI more efficient, especially with respect to design and allocation of the samples. The updated samples of items, outlets, and areas were based upon data from the Survey of Consumer Expenditures for the years 1982, 1983, and 1984; the 1980 Census of Population; and the ongoing Point-of-Purchase Survey, which, beginning with 1985, reflected the new item and area design. The new technique of rolling in the new area, item, and outlet samples significantly reduced the cost of introducing new samples. In addition, the housing survey was redesigned to represent optimally both owners and renters, which improved the estimation method for shelter costs.

The improvements introduced over the years have reflected not only the Bureau's own experience and research, but also the criticisms and investigations of outsiders. A major study was conducted during World War II by the President's Committee on the Cost of Living.[8]

The House Committee on Education and Labor conducted a detailed examination of the index in 1951.[9]

A decade later, a study was made by the Price Statistics Review Committee, which was appointed by the National Bureau of Economic Research, at the request of the Office of Statistical Standards of the Bureau of the Budget, to review all Government price statistics.[10]

A continuing flow of articles in professional journals and books has also contributed to the assessment of the CPI's quality and of the ways in which it might be improved.[11]

Improvements made to the CPI since 1966 are summarized in the chronology on the following page. The next major revision of the CPI will take effect in January 1998. Consumer Expenditure Survey data from 1993-95 will be used to calulated a new expenditure weight for each item strata category in every CPI index area. These new market baskets—new geographic area samples, new item structure, and new expenditure weights—will take effect with the index for January 1998. At the same time, many of the samples underlying the CPI will be replaced. These samples include geographic areas, items selected for pricing, and outlets in which items are priced.

[3] *Interim Adjustment of Consumers' Price Index*, Bulletin 1039, Bureau of Labor Statistics, 1951

[4] *Consumer Prices in the United States*, 1953-58, Bulletin 1256, Bureau of Labor Statistics, 1959.

[5] *The Consumer Prices Index: History and Techniques*, Bulletin 1517, Bureau of Labor Statistics, 1966.

[6] *The Consumer Price Index: Concepts and Content Over the Years*, Report 517, Bureau of Labor Statistics, 1978.

[7] "Changing the Homeownership Component of the Consumer Price Index to Rental Equivalence," *CPI Detailed Report*, January 1983, PP-7-13.

[8] *Report of The President's Committee on the Cost of Living* (Washington, Office of Economic Stabilization, 1945.

[9] *Consumers' Price Index, Report of a Special Subcommittee of the Committee on Education and Labor*, U.S. Congress, House of Representatives, 82/1, Subcommittee Report No. 2 (Washington, U.S. Government Printing Office, 1951).

[10] *Government Price Statistics, Hearing before the subcommittee on Economic Statistics, U.S. Congress*, Joint Economic Committee, 871. Part 1 (Washington, U.S. Government Printing Office, Jan. 24, 1961).

[11] For a list of published papers on the CPI, see the technical references at the end of this chapter.

Improvements to the Consumer Price Index: Chronology, 1966 to 1996

Change	Date Implemented	Description
New construction	1966	Rent samples augmented with units built after 1960.
Quality adjustment of new changeovers	1967	New automobile prices adjusted for quality differences after model automobile prices.
Sample rotation	1981	Introduced a systematic replacement of outlets between major revisions.
Rental equivalence	1983	Changed homeowners' component from cost of purchase to value of rental services for CPI-U
Return from sale price imputation	1984	Introduced procedure to eliminate downward bias for items discontinued by outlets that went out of index with discounted prices.
Rental equivalence	1985	Changed CPI-W homeowners' component to value of services.
Enhanced seasonal products methodology	1987	Enhanced methodology used for seasonal items by expanding the number of price quotations to select products from alternate seasons and eliminate under-representation of such items.
Quality adjustment of used car prices	1987	Prices of used cars adjusted for differences in quality after model changeovers.
Aging bias correction	1988	Rental values adjusted for aging of the housing stock.
Imputation procedures for new cars and trucks	1989	Price changes for noncomparable new models are imputed using only the constant-quality price changes for comparable model changeovers.
Quality adjustment of apparel prices	1991	Regression models used to adjust apparel prices for changes in quality when new clothing lines are introduced, and eliminate bias due to linking product substitutions into the CPI
Discount air fares	1991	Substitution rules modified to expand pricing of discount airline fares.
Sample augmentation	1992	Increase in the number of outlets from which prices are collected to replace sample lost through sample attrition.
New models imputation	1992	Refined imputation methods used when introducing products into the CPI.
Hotels and motels	1992	Samples for hotels and motels quadrupled to reduce variances related to seasonal pricing.
Seasonal adjustment	1994	Procedures for seasonal adjustment revised to eliminate residual seasonality effects.
Quality adjustment for gasoline	1994	Treat "reformulated" gasoline as a quality change and adjust the price to reflect quality difference. Impact of the change estimated.
Generic drugs	1995	Introduced new procedures that allow generic drugs to be priced when a brand drug loses its patent.
Food-at-home base period prices	1995	Introduced seasoning procedures to eliminate upward bias in setting of base period prices of newly initiated items.
Rental equivalence	1995	Modified imputation of homeowners' implicit rent to eliminate the upward drift property of the current estimator.
Composite estimator used in housing	1995	Replaced current composite estimator with a 6-month chain estimator. Under reporting of 1-month rent changes had resulted in missing price change in residential rent and homeowners' equivalent rent. Old estimator also produced higher variances.
Commodities and services base period prices	1996	Extended food-at-home seasoning procedures to remainder of commodities and service series. Base period prices left unchanged in most noncomparable substitutions.

Concepts

Several key concepts dictate how the Consumer Price Index should be calculated.

Prices and living costs

The CPI is based on a sample of prices of food, clothing, shelter and fuels, transportation, medical services, and other goods and services that people buy for day-to-day living. Price change is measured by repricing essentially the same market basket of goods and services at regular intervals and comparing aggregate costs with the costs of the same market basket in a selected base period.

A unifying framework for dealing with practical questions that arise in construction of the CPI is provided by the concept of the cost-of-living (COL) index.[12]

As it pertains to the CPI, the COL index for the current month is based on the answer to the following question: What is the cost, at this month's market prices, of achieving the standard of living actually attained in the base period?" This cost is a hypothetical expenditure—the lowest expenditure level necessary at this month's prices to achieve the base-period's living standard. The ratio of this hypothetical cost to the actual cost of the base-period consumption basket in the base period is the COL index.[13]

The COL index is a measure of price change (it compares current-period and base-period prices). However, the concept is difficult to implement because it holds the standard of living constant, and the living standard must be estimated in some way.

The CPI uses a fixed market basket to hold the base-period living standard constant. The CPI equals the ratio of the cost of the base-period basket at this month's prices to the actual cost of the base-period basket in the base period. It is an index of price change only and does not reflect changes in buying patterns that consumers probably would make to adjust to relative price changes. The formula used for calculating the CPI is the one known in price index literature as the Laspeyres index. (See part II.) The CPI provides an approximation to a COL index as a measure of consumption costs. It is sometimes said that the CPI's Laspeyres formula provides an "upper bound" on the COL index.

Note that both the CPI and the COL index that were defined above measure changes in expenditures. Neither one measures the change in income required to maintain the base-period living standard. For this reason, neither the COL index nor the CPI are affected by changes in income taxes, but both will include the effects of changes in sales taxes and other indirect taxes.

For certain purposes, one might want to define price indexes to include, rather than exclude, income taxes. One could develop either a COL index or a Laspeyres index along these lines. Such indexes would provide an answer to a different question from the one for which the present CPI is relevant, and would be appropriate for different uses. For a research measure of a consumption index inclusive of income taxes and Social Security contributions, see Gillingham and Greenlees.[14]

Sampling

Since it is not practical to obtain prices for all consumer transactions in the United States, the CPI is estimated from a set of samples. These samples, which use statistical procedures to make the CPI representative of the prices paid for all goods and services purchased by consumers in all urban areas of the United States, are:

- Urban areas selected from all U.S. urban areas,
- Consumer units within each selected urban area,
- Outlets from which these consumer units pur chased goods and services,
- Specific, unique items—goods and services—purchased by these consumer units, and
- Housing units in each urban area for the shelter component of the CPI.

Weights and relative importance

The weight of an item in the CPI is derived from the expenditure on that item as estimated by the Consumer Expenditure Survey. This survey provides data on the average expenditure on selected items, such as white bread, gasoline, and so on, that were purchased by the index population during the survey period. In a fixed-weight index such as the CPI, the implicit quantity of any item used in calculating the index remains the same from month to month.

A related concept is the relative importance of an item. The relative importance shows the share of total expenditure that would occur if quantities consumed were unaffected by changes in relative prices and actually remained constant. Although the implicit quantity weights remain fixed, the relative importance changes over time, reflecting the effect of price changes. Items registering a greater than average price increase (or smaller decrease) become relatively more important. Conversely, items registering a smaller than average price increase (or larger decrease) become relatively less important. Thus, the relative importance of medical care in the index for all urban consumers, which was 6.0 percent in December 1982, was 6.9 percent in December 1986. During the same period, the relative

[12] On the use of a cost-of-Living index as a conceptual framework for practical decision making in putting together a price index, see Robert Gillingham, "A Conceptual Framework for the revised Consumer Price Index," Proceedings of the Business and Economic Statistics Section, American Statistical Association, 1974, pp. 46-52.

[13] For more information on the cost-of-living concept, see the technical references at the end of this chapter.

[14] Gillingham, Robert F. and Greenlees, John. "The Impact of Direct Taxes on the Cost of Living, " *Journal of Political Economy*, 95(4), August 1987.

importance of energy fell from 12.4 percent to 8.9 percent. The published data on relative importance are often used to answer such questions as: What was the direct effect on the overall CPI of a particular price change (e.g., gasoline prices) for a particular period? (See appendix 2.)

Owners' equivalent rent

The concept of owners' equivalent rent used to measure homeowner shelter costs was introduced in the CPI-U in January 1983 and in the CPI-W in January 1985. The owners equivalent rent index measures the change in the cost of renting housing services equivalent to those services provided by owner-occupied housing.

Prior to the introduction of owners' equivalent rent, homeowners' shelter costs in the CPI were represented by five elements: (1) House prices, (2) mortgage interest costs, (3) property taxes, (4) homeowner insurance charges, and (5) maintenance and repair costs. These constitute the major costs associated with purchasing and maintaining the physical asset of a house.

This "asset price" approach to homeowner costs failed to distinguish the investment aspect of owning a home from the consumption aspect. The basic concept of the CPI is as a measure of the average change in the prices paid by consumers for consumption goods and services. Investment purchases, such as stocks and bonds, are conceptually out of the scope of the index and are excluded. A house is not consumed at the time of purchase. It is a long-lived asset (investment), but it also provides the owner with a flow of shelter services over time. Thus, it is the cost of this shelter service provided by the asset that is the conceptually appropriate element of the CPI.

To implement the new concept, the old homeownership component was replaced with two items: (1) Owners' equivalent rent; and (2) household insurance, which contains those parts of homeowners' insurance that do not insure the structure.

In addition, the previous maintenance and repairs component was made a new component covering both renters' expenses and owners' expenses—exclusive of those estimated to be part of owners' equivalent rent. Also, the weight for household appliances was reduced to remove those expenses in homeowners' costs for appliances included with the house.

Scope and Calculation

Prices for the goods and services used to calculate the index are collected in 85 urban areas throughout the country and from about 21,000 retail and service establishments-supermarkets, department stores, gasoline stations, hospitals, etc. In addition, data on rents are collected from about 40,000 landlords or tenants and 20,000 owner occupants are asked about their housing units.

Food, fuels, and a few other items are priced monthly in all 85 locations. Prices of most other goods and services are collected monthly in the five largest urban areas and bimonthly in the remaining areas. Trained BLS field representatives collect all price information through visits or telephone calls.

The index is the average of the price changes for all the various items within each area. The averaging process uses weights which represent the importance of the items in the spending pattern of the appropriate population group in that area. Separate indexes are compiled for: the urban United States, 4 regions, 4 size classes, 13 groups cross-classified by region and population size, and 29 local metropolitan statistical areas.

Movements of the indexes from one month to another are usually expressed as percent changes rather than changes in index points. Index point changes are affected by the level of the index in relation to its base period while percent changes are not. The example in the tabulation illustrates the computation of index points and percent changes:

Index point change

CPI	136.0
Less CPI for previous period	129.9
Equals index point change	6.1

Percent change

Index point difference	6.1
Divided by the previous index	129.9
Equals	0.047
Results multiplied by 100.0.	0.047 x 100
Equals percent change	4.7

Percent changes for periods that are less than 1 year can be expressed as annual rates and are computed according to the standard formula for compound growth rates. These data indicate what the percent change would be if the average rate for the period was maintained for a 12-month period.

Uses

Almost all Americans are affected by the Consumer Price Index because of the many ways that it is used. Three major uses are:

As an economic indicator. As the most widely used measure of inflation, the CPI is an indicator of important economic trends. The President, Congress, and the Federal Reserve Board use trends in the CPI to aid in formulating fiscal and monetary policies. In addition, business executives, labor leaders, and other private citizens use the index as a guide in making economic decisions.

As a deflator of other economic series. The CPI and its com-

ponents are used to adjust other economic series for price changes and to translate these series into inflation-free dollars. Examples of series adjusted by the CPI include retail sales, hourly and weekly earnings, and components of the gross domestic product.

As a means of adjusting income payments. Almost 3 million workers are covered by collective bargaining agreements which tie wages to the CPI. The index affects the income of more than 70 million persons as a result of statutory action: 43.1 million Social Security beneficiaries, about 3.9 million military and Federal Civil Service retirees and survivors, and about 22.6 million food stamp recipients. Changes in the CPI also affect the cost of lunches for the 24.2 million children who eat lunch at school. Some private firms and individuals use the index to keep rents, royalties, alimony, and child support payments in line with changing prices. Finally, since 1985, the CPI has been used to adjust the Federal income tax structure to prevent inflation-induced increases in tax rates.

Analysis and Presentation

CPI data are issued initially in a news release about 2 weeks following the reference month. Summary tables are sent to persons on the mailing list at that time. The data are also available on the Internet, in recorded telephone messages, and by fax. The *CPI Detailed Report*, available about 3 weeks after the initial release, provides detailed indexes and a monthly analysis of U.S. price movements. *The Monthly Labor Review* also contains much of the CPI data each month and provides regular analyses of recent price movements as well as of long-term trends.

Seasonally adjusted data—primarily of use for current economic analysis—are presented in addition to the unadjusted data.

Limitations of the Index

CPI users should understand that the CPI may not be applicable to all questions about price movements for all population groups. For example, the indexes represent the average movement of prices for the U.S. urban population and, thus, may not be appropriate for use by non-urban residents. Also, the CPI does not provide data separately for the rate of inflation experienced by any demographic subgroup of the population, such as the elderly, which may purchase different consumer items and face different rates of inflation. And, as mentioned earlier, the CPI is an index of price change, not a cost-of-living index.

In addition, the indexes cannot be used to determine relative living costs. The change in an index for individual geographic area index measures how much prices have changed in that particular area. It does not show whether prices or living costs are higher or lower in that area relative to another.

A further limitation is that the CPI is not a complete measure of price change. Because the index is estimated from a sample of consumer purchases, the results may deviate slightly from those which would be obtained if all consumer transactions were covered. These estimating or sampling errors are statistical limitations of the index

A different kind of error in the CPI can occur when a respondent provides BLS field representatives with inaccurate or incomplete information. The Bureau attempts to minimize these errors by obtaining prices by personal observation wherever possible, and by correcting errors immediately upon discovery. The field representatives, technicians, and commodity specialists who collect, process, and analyze the data are trained to watch for deviations in reported prices which might be due to errors. Also, an independent audit staff conducts a systematic evaluation of all CPI collection and processing activities. The goal is to develop long-term quality improvement in all aspects of the index calculation.

A fuller discussion of the varieties and sources of possible error in the index is presented in part III of this chapter, "Precision of Estimates".

The construction of the Consumer Price Index is based on a series of samples—and on estimation procedures described below.

Part II. Construction of the Index

Definition of the Index

The CPI is defined as a fixed-quantity price index, that is, a measure of the price change in a fixed market basket of consumption goods and services of constant quantity and quality bought on average by urban consumers, either for all urban consumers (CPI-U) or for urban wage earners and clerical workers (CPI-W). It is a ratio of the costs of purchasing a set of items (i) of constant quality and constant quantity in two different time periods. We denote the index by It,0, where t is the comparison period for which a new index number is to be calculated and 0, the reference period:

$$I_{t,0} = \frac{\sum_i P_{it} Q_{ib}}{\sum_i P_{i0} Q_{ib}} \times 100.0$$

where:

- P_{it} is the price for the ith item in comparison period t
- P_{i0} is the price for the ith item in reference period 0
- Q_{ib} is the quantity of the ith item consumed in the expenditure base period b.

When the expenditure base (b) and reference period (0) coincide, this becomes the Laspeyres price index formula. For the 1987 revision of the CPI, however, they did not coincide and the formula was modified. The expenditure data, Pib Qib, from the 1982-84 Consumer Expenditure Survey (described later in this chapter) were updated for relative price changes (Pip/Pib) to November or December 1986, the respective pivot periods, p, when they were introduced into the CPI. Expenditure data for index areas priced bimonthly in the odd-numbered months were updated to November 1986. Expenditure data for the index areas priced monthly or bimonthly in the even-numbered months were updated to December 1986. Price relatives from the midpoint (June 1983) of the Consumer Expenditure Survey (CE) collection period to November or December 1986 were calculated from the appropriate detailed indexes for the period. The price relatives were applied to the corresponding expenditure data, thereby updating the expenditures to the end of 1986. Continuity with the pre-1987 version of the CPI was maintained in the published version by modifying the above formula to:

$$I_{t,0} = \frac{\sum_i P_{it} Q_{ib}}{\sum_i P_{i0} Q_{ib}} \times I_{p,1967}$$

where Ip,1967 is the 1967-based value of the CPI for the pivot month, November or December 1986.

NOTE: The base period for the expenditure weights, 1982-84, should not be confused with the numeric reference base period, although both are currently 1982-84=100. A new numeric reference base period for the index was established effective with the release of the index for January 1988, when the index was rebased to 1982-84=100. The previous numeric reference base was 1967=100. Such changes in numeric bases are achieved by dividing all indexes in a series by the appropriate average index on the old base for the time period of the new base.

Index Calculation

For sampling and index computation purposes, the set of all retail consumer expenditures by the target population for a given index area has been subdivided into 207 classes of similar items (see appendix 4) called item strata. The item strata are mutually exclusive and exhaustive of all consumer expenditures. They are defined identically for both index populations.

There are currently 88 areas (see appendix 3) which combine to yield the 41 basic areas—29 self-representing areas plus 12 region- and population-size class cross-classifications for non-self-representing areas. Thus there are 8487 (207 times 41) basic CPI strata. Each month BLS calculates indexes for these basic strata and aggregates these indexes to higher area and item groups.

The CPI is computed by a chaining process in which the index for the previous month, Ihzt-1,0, for each basic area (h) and item stratum (z) is multiplied by an estimate of the relative change in price from the previous month to the current month to provide the current-month index for that area and item stratum:

$$Ihzt,o = Ihzt\text{-}1,0 \times Rhzt,t\text{-}1$$

where Rhzt,t-1 is an estimate of the one-period price change in the hth basic area for the zth item stratum. Basic area

and item stratum indexes are then aggregated using aggregation weights to form desired aggregate area and item indexes, $I_{HZt,0}$:

$$I_{HZt,0} = \frac{\sum_{h \varepsilon H} \sum_{z \varepsilon Z} A_{hz} I_{hzt,0}}{A_{HZ}}$$

where A_{hz} is the aggregation weight for the hth basic area and zth item stratum A_{HZ} is the aggregate weight for the aggregate area and item index, and where H and Z represent the set of h areas and z items, respectively.

The U.S. city average all items CPI is computed by aggregating all basic component area and item aggregate indexes:

$$I_{ust,0} = \frac{\sum_{h \varepsilon US} \sum_{z \varepsilon US} A_{hz} I_{hzt,0}}{A_{US}} \quad \text{for all } h, z$$

Aggregation weights for a given area H and item stratum Z combination are calculated as the expenditures for the pivot period divided by the corresponding index in the pivot period, that is,

$$A_{hz} = \sum_{h \varepsilon H} \sum_{z \varepsilon Z} P_{hzp} Q_{hzb} / I_{hzp,r}$$

where p is November or December 1986 and r is the numerical base, currently the same as b (the weight-base), equal to the average of 1982-84. The pivot period is the time when revised CPI is attached to the previous CPI to create a continuous series.

The formula for the aggregation weights of high level item and area indexes is:

$$A_{HZ} = \sum_{h} \sum_{z} P_{hzp} Q_{hzP} / I_{HZp}$$

The aggregation weight for higher level indexes is not, in general, equal to the sum of the aggregation weights of its elements.

The computation of the index using one-period price relatives based on identical item specifications in adjacent periods allows the requisite flexibility to update the samples of outlets and specifications to reflect an updated distribution of purchases within an item stratum within a local area. The form of the estimator for a one-period price change, Rhzt,t-1, depends on the procedures used for selecting the samples of outlets and items and are designed to provide approximate unbiased estimators of price changes. When the samples are selected with each unit having a probability proportional to quantity, the estimator of $R_{hzt,t-1}$ is:

$$R_{hzt,t-1} = \frac{\sum_{z \varepsilon z} W_{hz} P_{hit}}{\sum_{z \varepsilon z} W_{hz} P_{hi,t-1}}$$

This is the ratio of the summation of weighted prices, where the weights (W_{hz}) reflect the probability of selection of the item being priced and a non-interview adjustment. Non-interview adjustment is a statistical procedure designed to adjust for nonresponse. This form of the estimator is used in the rent and owners' equivalent rent item strata.

When the samples of outlets and items are selected 'with probability proportional to expenditure, the estimator of $R_{hzt,t-1}$ is:

$$R_{hzt,t-1} = \frac{\sum_{z \varepsilon z} W_{hz} P_{hz,t} / P_{hia}}{\sum_{z \varepsilon z} W_{hz} P_{hz,t-1} / P_{hia}}$$

where P_{hza} is an estimate of the price of the selected item in period a, corresponding to the expenditures used in outlet sampling.

This is the ratio of the summation of weighted price ratios, where the weights reflect the probability of selection of the item being priced and non-interview adjustments. It is used for all commodity and service item strata.

Thus, construction of the CPI is a twofold estimation process. First, the aggregation weights. Ahz, must be estimated. These estimates are derived from the Consumer Expenditure Survey as explained in the next section. Second, the one-period price changes, $R_{hzt,t-1}$, must be estimated for each pricing period. The methodology for estimating price changes is explained in later sections.

Estimation of expenditure-population weights

The aggregation weights for each of the index areas, for each item stratum, and for both the CPI-U and CPI-W populations, require an estimate of expenditure that serves as its weight in the market basket of goods and services. When divided by their respective index for the pivot period, the expenditures weight becomes the aggregate weight. Each expenditure weight is the product of the mean annual expenditure per consumer unit (estimated from the 1982-84 CE Survey) and the number of consumer units (obtained from a special tabulation of the 1980 census).

Calculation of mean expenditures consists of three steps: (1) Preliminary estimation of expenditures and their relative importance; (2) final estimation of relative importance data, using a composite estimation procedure; (3) final esti-

mation of mean expenditures using a raking process on the mean expenditures derived from the relative importance data determined in (2). The design criterion for this estimation procedure is to minimize the average mean square error of the relative importance of the expenditure weights for the index areas. The relative importance is a stratum's share (usually expressed as a percent) of the total expenditure for consumption.

Preliminary mean expenditures and relative importance. Preliminary mean expenditures and their relative importance were calculated for each area-item stratum and expenditure class, for each population, index area, replicate,[15] and major geographic area from each survey source—the CE Interview or Diary Surveys. Composite estimation reduced the variance of these preliminary relative importance data by averaging them with the relative importance of their major geographic areas. There were eight major geographic areas (index area aggregates), each consisting of either the self-representing or non-self-representing index areas[16] in each Census region. The mean expenditures are estimated using information from the 1982, 1983, and 1984 CE Surveys. They are the simple weighted averages (overall consumer units in the population desired) of the expenditures for the particular stratum or expenditure class. The consumer unit weights are those described in the section on the CE Survey.

The relative importance of stratum or expenditure class is the proportion of total consumption expenditures that the consumer units in a population index area (or major area), and replicate spend for that structure or class.

Composite estimation. Composite estimation is a method used to decrease the mean square error (MSE) of the relative importance of area-item strata by using data from the corresponding major areas. It is implemented separately for data from the Diary and Interview Surveys. The composite estimated relative importance of a particular area-item stratum, or expenditure class is a weighted average of the two preliminary relative importance estimators, RI, one for the index area and major area.

To calculate the composite estimator, let the relative importance for the particular item stratum (or EC) at the index area and one for the major area be R_{Ih} and R_{Im} respectively. The initial composite estimator, ICRI, is a weighted average of the PSU level and major area. It is calculated as:

$$ICRI = B_h \cdot RI_m + (1 - B_h) \cdot RI_h$$

where:

$$B_h = \frac{\text{Var } RI_h - \text{Cov } RI_h}{\text{ESD } RI_h}$$

where:

Var R_{Ih}, is the estimated variance of the relative importance for index area h,

Cov R_{Ih}, is the estimated covariance of the relative importance RI_h and RI_m and

ESD R_{Ih}, is the estimated expected squared difference of the relative importance R_{Ih} and RIm.

If B_h <0, then B_h is set equal to O. If B_h > 1.0, then B_h is set equal to 1.0.

The above composite estimation procedure defines a "shrinkage" estimator. A further refinement to defining the final composite estimator, CRI, limits the shrinkage. Let SDRIh be the square root of the Var RIh. The final composite estimator is calculated as:

where $K_{m,mg}$ is a limiting factor defined for each major area, m, and major item group, mg. The parameter, $K_{m,mg}$ was determined by testing alternative values on the 1972-73 CE data. The value which maximized the reduction of the MSE and minimized the change for expenditure estimates at the major group level was selected.

Raking. To reestablish data consistency between item strata and EC levels and to reflect the special consumer unit weights, an iterative ratio estimation procedure (raking) was performed. That is, the sum of the expenditures for all item strata within an EC for an index area was forced to equal the total expenditures for the EC in the index area. The sum of the expenditures for a specific item stratum across all index areas in a major area was forced to equal the major area estimate of average expenditures for the item stratum times the special consumer unit weights. Alter composite estimation of the relative importance, initial total expenditure estimates, TE, for each item stratum and expenditure class, population, index area, and replicate for the Diary and Interview Surveys were calculated as:

where ES_z is the sum of preliminary mean expenditures across all item strata for index area z. Similarly, total expenditure estimates were calculated for each item stratum and EC at the major area level.

Expenditures from the Diary and Interview Surveys were then integrated. Each item has as its expenditure source either the Diary or Interview Survey. All of the processing

[15] A single selection of entry level items and outlets for all item strata assigned to a primary sampling unit is called a replicate. For further information, see the section on sampling.

[16] Self representing and non-representing index areas are defined in the section on sample and publications areas. Self representing and non-representing index areas are defined in the section on sample and publications areas.

activities described above were performed separately by survey. After composite estimation and rakingwere completed, expenditures from the two surveys were combined to form the set of total expenditures.

The raked, composite estimated expenditures were converted to aggregation weights in a two-step process. First, item strata expenditures were updated from June 1983, (the midpoint of the 1982-84 CE period) to December 1986 (the pivot month) for the index area. (November 1986 for bimonthly index areas published in odd months.) Each expenditure weight was multiplied by a long-term price change from the CPI for the time interval:

$$E_{hz8612} = E_{hz8306} \frac{I_{hz8612}}{I_{hz8306}}$$

where:

E_{hz8612}	is the updated expenditure for the item strata in the h_{th} area for December 1986,
E_{hz8306}	is the raked composite estimated expenditure for the z_{th} item strata in the hth area for June 1983
I_{hz8612}	is the CPI index for the z^{th} item strata for the hth area for December 1986, and
I_{hz8306}	is the CPI index for the z^{th} item strata for the h_{th} area for June 1983.

Updated item strata expenditures were summed to arrive at the required updated expenditures for aggregate items and areas. In the second step, the updated expenditure weights were divided by a corresponding index for the pivot period, yielding the aggregation weight:

$$A_{hz} = \frac{E_{hz8612}}{I_{hz8612}}$$

Special expenditure-weight procedures. As a result of the 1987 revision, the cost-weight definitions for new and used vehicles and for medical care item strata were changed significantly. In the former CPI, all expenditures for vehicles were valued at the net transaction price—the negotiated price less any trade-in value. In the revised CPI, the treatment of trade-ins and outright sales of used vehicles was changed. Trade-ins at their market value continue to be netted out of the price of used vehicles. However, the total market value of trade-ins on new vehicles is now netted from used vehicle purchases, rather than new vehicles. Sales of vehicles from one consumer to another are netted against the corresponding purchase because the transaction is really an intrapopulation exchange of wealth; there is no net change in the consumption of used vehicles. As a result of these changes, the expenditure weight for used vehicles represents: (1) Purchases of used vehicles that are entering the consumer sector for the first time (that is, purchases from private industry, including auto rental companies, and from public agencies; as well as purchases of used vehicles that enter the U.S. market from abroad as used vehicles); and (2) the difference between, (a) the price a consumer receives when he sells (or trades-in) a vehicle to an auto dealer, and (b) the price another consumer pays when he buys the vehicle from that (or a different) auto dealer.

In medical care, the way health insurance premiums were represented changed in the expenditures weights. While this change has no effect on the final index result and is mathematically equivalent to the former procedure, the new structure provides a clearer picture of the role of health insurance in the CPI. In both the former and current CPI, the health insurance weight is only expenditures by consumers for premiums; it does not include employers' contributions, Just as other medical care expenditures are out-of-pocket payments by consumers. Insurance premiums can be viewed as purchasing two things: (1) the medical care for which benefits are paid, and (2) the services of the insurance carrier in administering the policy and bearing risk. This second element has been labeled retained earnings and refers to operating costs and any profit of the insurance carrier.

In the former CPI, within health insurance, the insurance expenditure weight was broken into item strata for each type of benefit and for the retained earnings associated with each type of benefit. The price movement for a health insurance benefit stratum (for example, insurance-paid hospital rooms) was the same as the price movement for the corresponding medical items in the CPI (hospital rooms). The price movement for a retained earnings stratum was the combination of price change for the relevant medical care items and an estimate of changes in retained earnings as a proportion of premiums.

In the current CPI, using the same example, instead of using the price change for both the hospital room index and the hospital-room-paid-by-insurance index, the expenditures for the two types of hospital payment are combined into a single index. The cost-weight for each medical care item is the combination of direct out-of-pocket expense for the item and indirect out-of-pocket expense for the item paid from consumer-purchased health insurance. The health insurance cost weight is the sum of all retained earnings

Annual and semiannual average index estimation

CPI values use 12 successive months of CPI values as:

$$I_{12av} = \sum_{t=1}^{12} I_{t,0} / 12$$

Semiannual average indexes are computed for the first half of the year (January-June) and for the second half of the

year (July-December) using 6 successive months of CPI values as:

$$I_{6\,av} = \sum_{t=1}^{6} I_{t,0} / 6$$

where the value of each monthly index is real or interpolated, depending on availability.[17]

For bimonthly indexes, the intermediate indexes are calculated using a geometric mean of the values in the months adjacent to the one being estimated.

Sampling: Areas, Items, and Outlets

The 1987 CPI area sample defined 91 geographic areas from which 94 PSUs were selected. In 1988 due to budget constraints, 6 PSUs were dropped from the design leaving 88 PSUs in 85 areas. The following sections describe the sample as it was designed for the 1987 revision. Where changes have occurred due to the 1988 budget constraints, the resulting numbers will be noted in parenthesis next to the original numbers. See appendix 3 for more details.

Area sample

Pricing for the CPI is conducted in 94 (88) primary sampling units (PSU's) in 91 (85) geographic areas. (The New York area has three PSU's and the Los Angeles area has two PSU's.) The area design[18] and sampling are summarized as follows: The urban was divided into 1,088 PSU's. Except in New England where they are groups of cities and towns, a PSU is a county or a group of contiguous counties. The basis of the PSU definition was the geographic areas defined by the Bureau of the Census for the Current Population Survey in 1980 with population estimates from the 1980 census. Each Consolidated Metropolitan Statistical Area (CMSA) or Metropolitan Statistical Area (MSA) as defined by the Office of Management and Budget in 1983 is a PSU. BLS grouped the remaining non-MSA counties containing any urban population to form PSU's. Rural areas of the non-MSA counties were excluded. (See appendix 3.)

The urban U.S. was partitioned into 91 geographic strata. The strata were sets of similar PSU's combined according to the following characteristics, which were found to be highly correlated with price change:

a. Region, population size, MSA vs. non-MSA,
b. Mean interest and dividend income per housing unit,
c. Mean wage and salary income per housing unit,
d. Percent of housing units heated by electricity,
e. Percent of housing units heated by fuel oil,
f. Percent black, and
g. Percent retired.

The Census Bureau defined the four regions: the Northeast, the Midwest, the South, and the West. These characteristics along with preference factors for each characteristic, the number of strata, and the limits on strata size were the input to a program which used a modified Friedman-Rubin clustering algorithm to determine the stratification. This algorithm attempts to identify the best partition of sampling units into strata by minimizing the between PSU variance. "

This area design resulted in 31 strata with one pricing area per stratum (self-representing PSU's) and 60 non-self-representing strata. (The three New York PSU's and two Los Angeles PSU's are also self-representing.) One sample PSU was selected from the PSU's in each non-self-representing stratum. A controlled selection program was used to insure that the sample areas were distributed geographically across the United States and to increase the overlap between the old area sample and the new area sample.

Since 39 of the 94 PSU's selected were new to the CPI the new area sample for the CPI was introduced over a 2-year period. Twenty new PSU's were initiated during 1986 and have been used in CPI index calculations since January 1987. Sixteen of the remaining new PSU's were initiated during 1987, and the final three were initiated in 1988. Each of the 19 old PSU's continued to be priced until the new PSU which corresponded to it was initiated and linked into the index.

The CPI area design defines 37 (36) publication areas, that is, areas for which a CPI is published. Twenty-seven of the self-representing areas were defined as publication areas. Eight (seven—the B size class in the West is no longer published due to budgetary constraints) additional publication areas were defined by crossing the two city-size classes (non-self-representing MSA areas) by the four Census regions. The non-MSA areas in the Midwest and South were also defined as publication areas. Each of these region-by-size publication areas has 4, 6, or 10 strata. Only two strata were defined in the non-MSA areas in the West and Northeast, which made them ineligible for publication. Indexes are also published for the U.S. total as well as for region and city-size class totals.

Indexes for the U.S. total, the 10 (9) region-by-size class areas, and the 5 largest local areas are published monthly. Indexes for the 10 next largest areas are published bimonthly, and indexes for the smaller self-representing areas are published only as semiannual and annual averages.

Every 2 months the CPI prices 134 (126) replicates. One replicate is approximately 1,100 price quotes for commodities and services and 390 housing units for shelter. The allocation of the replicates is proportional to the population rep-

[17]To be published, a semiannual average must have at least two noninterpolated index values with sufficient samples. An annual average must have at least four noninterpolated index values with sufficient samples.

[18]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, pp. 118-123.

resented by the PSU, with at least one replicate assigned to each PSU. The actual allocation of replicates to PSUs is provided in the next section.

Item and outlet samples: Commodities and services other than shelter

Item structure and sampling. The CPI item structure bas four levels of classification. The 7 major groups are made up of 69 expenditure classes (EC's), which in turn are divided into 207 item strata. Within each item stratum, one or more substrata, called entry level items (ELI's), are defined. There are 364 entry level items. (See appendix 4 for a complete list of consumer expenditure classes, item strata, and ELI's.) The major groups are primarily for emphasis in publication. The Expenditure categories are needed to smooth the weights of the item strata during composite estimation. (See above.) The CPI uses the item strata at the CPI area level to calculate the index each month. The ELI's are the ultimate sampling units for items as selected by the BLS national office. They are the level of item definition at which the data collectors begin item sampling within each sample outlet.

To enable the CPI to reflect changes in the marketplace, new item and outlet samples are selected each year for 20 percent of the PSU's on a rotating basis. Each year, four regional item-universes are tabulated from the two most recent years of Consumer Expenditure Survey data. An independent sample of ELI's is selected for each item stratum for each PSU-replicate scheduled for rotation that year from the corresponding regional item universe. For the complete 5-year cycle, 134 samples of ELI's are selected nationally. Within each region, each ELI sample uses a systematic sampling procedure, in which each ELI has a probability of selection proportional to the CPI-U population expenditures for the ELI within its stratum. Selection of outlet samples is described in the following section.

Item and outlet sample design. The CPI uses two separate sample designs, one for rent and owners' equivalent rent and one for all other commodities and services. The methodology employed to determine the commodities and services item and outlet sample design is presented here. Those for the rent and owners' equivalent rent components are described later.

For the development of the sample design, all commodities and services item strata were grouped into eight major groups:

Food and beverages	Transportation
Fuels and utilities	Medical care
Household services and furnishings	Entertainment
	Other commodities and services
Apparel and upkeep	

The objective of the sample design methodology was to determine, by major group, the number of ELI's to be sampled and the number of outlets to be selected per PSU replicate. There were four major activities in the design project. First, a variance function projected the variance of price change as a function of the above variables for the commodity and service components. Second, a cost function modeled the total annual cost of the commodities and services components of the CPI. Third, values for all coefficients of the two functions were estimated (including estimates of outlet overlap). Fourth, nonlinear programming techniques were used to determine approximately optimal sizes for the item and outlet samples to minimize the CPI variance under varying assumptions of annual price change subject to cost constraints.

The variance and cost functions for the CPI were modeled for 10 PSU groups:

1. New York City
2. New York, New Jersey suburbs
3. Los Angeles City
4. Los Angeles suburbs
5. Chicago
6. Philadelphia
7. San Francisco, Detroit
8. Large self-representing PSU's
9. Small self-representing PSU's
10. Non-self-representing PSU's

A detailed discussion of the sample allocation methodology is provided in appendix 5.

The solution of the design problem yielded the following number of item strata selections per PSU replicate by Major item group:

CPI expenditure categories	*Item strata selections (ELI's)*
Food and beverages	73
Fuel and utilities	12
Household services and furnishings	66
Apparel and upkeep	47
Transportation	34
Medical care	18
Entertainment	27
Other commodities and services	21

The number of outlets selected for each Point-of-Purchase-Survey category (see below) for each major item group by PSU group is as follows:

OPS expenditure categories	*PSU group*									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Food and beverages	6	7	6	6	8	9	9	4	2	6
Fuel and utilities	7	8	4	4	7	8	6	4	2	6
Household services and										

furnishings	1	1	1	1	1	1	1	1	1	1
Apparel and upkeep	2	2	2	2	2	3	3	2	1	2
Transportation	2	4	3	3	3	4	4	3	1	3
Medical care	3	3	3	3	3	5	3	3	1	4
Entertainment	1	1	1	1	1	1	1	1	1	1
Other commodities and services.	1	1	1	1	1	2	1	1	1	1

The PSU groups are as defined earlier. With this allocation, outlets and quotes will be initiated each year under sample rotation. For ongoing pricing, there will be about 25,000 outlets visited each month, with prices collected for 95,000 items.

Outlet and price surveys. BLS field representatives collect prices monthly for food, energy items, and a small number of other commodity and service item strata in all PSU's. Prices are collected monthly for all commodity and service item strata in the five largest index areas (New York, Los Angeles, Chicago, Philadelphia, and San Francisco). Prices are collected bimonthly for the item strata not cited above in the remaining index areas. Each bi-monthly PSU was assigned to either the even- or odd-numbered months for pricing.

Point-of-Purchase Survey. The Bureau of the Census conducts the Continuing Point-of-Purchase Survey (CPOPS) for BLS. It furnishes current data on retail outlets from which urban households made purchases of defined groups of commodities and services. Data from the survey provide the sampling frame of outlets for food and most commodities and services to be priced in the CPI. CPOPS is the source of the outlet sampling frame for about 90 percent of the commodities and services items by expenditure weight. (See appendix 6 for a list of CPOPS categories. See appendix 7 for a complete list of items not covered by the CPOPS.)

The Point-of-Purchase Survey conducted in 1974 was the source of the outlet sampling frame in the 87 PSU's defined for the 1978 CPI revision. It was based on the 1970 census. From 1977 to 1984, the Continuing Point-of-Purchase Survey was conducted in approximately one-fifth of these PSU's on a rotating basis, so that the outlet sample for any PSU was never more than 5 years old. Since 1985, the survey has been based on the 1980 census and covers the 94 PSU's defined for the 1987 CPI.

Various methods have been tested to determine the sample of households to be interviewed in the survey. In 1974, a highly clustered sample of households was selected on the assumption that, if families tended to buy in the areas where they live, the outlets given as responses to the survey would also be clustered. In order to increase the expected chance of clustering the outlets, the household clusters were formed (where possible) around central business districts, shopping centers, and other retail centers. These large clusters were called secondary sampling units (SSU's). Within a cluster of tracts, a sample of Census enumeration districts (ED's) was selected, and, within the selected ED's, the sampled households were dispersed evenly. Five housing units were selected in each ED, and, since the desired sample size per cluster was 40 housing units, about 8 ED's were sampled from each cluster. In areas that issue permits for new construction, construction units were selected from the list of permits issued; in other areas, selection was made from area segments. This sampling approach was used, with minor modifications, in 1977 and 1978.

From 1979 through 1984, unclustered samples of households were selected for the survey. Since 1985, households have been selected on the basis of a noncompact clustering procedure which is modeled after the sample design for the Consumer Expenditure Survey. There are five sample frames: Unit, special place, area, block, and permit. The 1980 Census 100-Percent Edited Detail File is used as the source for all frames, except new construction in permit areas. For this frame, an unclustered sample of units is chosen from the permits issued since January 1980. For the unit, special place, area, and block frames, ED's are selected first. Then, a systematic sample of four or five units from each chosen ED is selected.

The Continuing Point-of-Purchase Survey is conducted annually over a period of 4 to 6 weeks, usually beginning in April, in approximately one-fifth of the PSU's in the CPI. The eligible population for the survey is the same as for the CE Survey: All civilian, non-institutional persons, including persons residing in boarding houses, housing facilities for students and workers, mobile home parks, permanent-type living quarters in hotels and motels, and staff residing in institutions. The interviews, conducted in selected housing units, consist of two parts. First, the interviewer elicits information on the demographic and socioeconomic characteristics of the household. This information is used to analyze the shopping patterns of various segments of the population. It is also used to determine how many consumer units reside in the housing unit and should be interviewed. A consumer unit (CU) consists of all members of a particular housing unit or other type of living quarters who are related by blood, marriage, adoption, or some other legal arrangement, such as foster children, or who are unrelated but financially dependent upon each other for major living expenses, such as housing or food.

In the second part of the interview, the respondent is asked whether or not the CU purchased categories of goods and services within a specified recall period. Commodities and services are grouped into sampling categories called POPS categories based on entry level items as defined in the CPI classification structure. Some POPS categories consist of only one ELI, while others consist of combinations of ELI's. ELI's are combined into a single POPS category when the commodities or services are generally sold in the same outlets. For example, POPS category 106, Meat and Poultry, con-

sists of eight beef ELI's, six pork ELI's, four ELI's for other meats, and three poultry ELI's. These ELI's are combined because an outlet that sells beef also tends to sell other meats.

Recall periods for POPS categories vary from 1 week to 5 years. The recall period for a specific POPS category is defined to produce a sufficient, but not excessive, number of outlets for sampling purposes. Since consumer units tend to purchase food items, tobacco products, and gasoline frequently throughout the year, a 1- or 2-week recall period is used. In contrast, CU's tend to purchase cars, hard-surface flooring, and funeral services infrequently; therefore, a 5-year recall period is assigned to these categories.

In the 1987 survey design, there are 170 POPS categories. Two different checklists of POPS categories are used by interviewers—each checklist is used in one-half of all sample households in the total nationwide sample. Each checklist consists of a subsample of 147 POPS categories. Most categories are included on both checklists. However, some of the short-recall-period categories are included on only one checklist. Subsampling on two checklists is used to control the expected number of responses received from a household and to minimize respondent burden. The combination of sample size and reference period for a given POPS category is designed to generate 6 to 12, not necessarily unique, outlets reported for a given PSU/POPS category. For each POPS category on the designated checklist, the respondent for a CU is asked whether purchases were made within the stated recall period and, if so, the names and locations of all places of purchase and the expenditure amounts.

From the results of the annual household survey, a new sample of outlets is selected for approximately one-fifth of the PSU's in the CPI. In the year following the survey, BLS initiates these new outlets, selects items for pricing from each, and replaces the former set of items in the CPI from each surveyed city with the new outlets and items.

Outlet sampling procedures. As indicated earlier, item samples and outlet samples are selected each year for 20 percent of the PSU's on a rotating basis. In self-representing PSU's, sample households for the POPS are divided into two or more independent groups by the first-stage order of selection, which defines two or more frames of outlets for outlet selection. The independent groups or replicates are needed for variance estimation. A single selection of ELI's and outlets for all item strata assigned to a PSU is called a replicate. For a given PSU, POPS category and replicate, the total expenditures reported for a given outlet are edited. If a purchase is reported for an outlet but the amount of expenditures is not reported, then, to ensure a chance of selection for the outlet, the mean expenditure for the PSU, replicate, and category is assigned. If large expenditures are reported for an outlet, then the amount is restricted to 20 percent of the total expenditure for the PSU, replicate, and category.

Outlet samples are selected independently for each PSU, replicate, and POPS category using a systematic sampling procedure. Each outlet on the frame has a probability of selection proportional to the amount of expenditures reported in the POPS. In each PSU replicate, all ELI's selected in the item sampling process are assigned for pricing to each sample outlet for the corresponding POPS category. When multiple selections of the sample outlet occur, a commensurate increase is made in the number of quotes priced for the outlet. The designated sample size for a given POPS category within each major item group for each replicate within a PSU group was presented in the section on item and outlet sample designs. The designated sample size is the number of outlet selections and not the number of unique outlets. The number of replicates by PSU group is presented below. There are 134 (126) replicates included in the CPI.

PSU group	*Number of replicates*
1. New York City.	4
2. New York suburbs, New Jersey suburbs	4
3. Los Angeles City	4
4. Los Angeles suburbs	2
5. Chicago.	4
6. Philadelphia	2
7. San Francisco, Detroit	4
8. Large self-representing PSU's	18
9. Small self-representing PSU's	32 (28)
10. Non-self-representing PSU's	60 (56)

Outlet sampling procedures for commodities and services not included in the POPS. Some commodity and service items were excluded from the POPS either because existing sampling frames were adequate or it was apparent that the POPS would not yield an adequate sampling frame. (See appendix 7.) For each non-POPS item, BLS either constructs the sampling frame or acquires it from another source. Each non-POPS item has its own sample design. The frames consist of all outlets providing the commodity or service in each sample area. A measure of size was associated with each outlet on the sampling frame. Ideally, this measure of size was the amount of revenue generated by the outlet for the item for the CPI-U population in the sample area. Whenever revenue was not available, an alternative measure of size, such as employment, number of customers, or sales volume, was substituted. All samples were selected using systematic sampling techniques with probability proportional to the measure of size.

The source of the sampling frame, the definition of the sampling unit, the measure of size employed, the final pricing unit, and the number of designated outlets and quotes for each non-POPS item are presented in appendix 7.

Augmentation. BLS monitors the sample of outlets for sufficiency on a regular basis. A sample of outlets for a given

PSU replicate and POPS category is sufficient if at least one-half of the designated outlets are still active; that is, prices are being collected in at least one-half of the originally specified sample. When a sample no longer meets the sufficiency criteria, BLS augments it, linking new outlets into the index. These augmentation samples are selected from the same universe as the original sample and have the same number of outlets designated. The newly initiated outlets and quotes are used in addition to any remaining outlets and quotes from the original sample. This process allows BLS to maintain a sufficiently large sample for each PSU, replicate and POPS category between the normal 5-year rotation cycle. Current BLS plans call for augmentation to occur on an annual basis.

Merging item and outlet samples. Since the item and outlet samples are selected separately, they must be merged before data collection. A concordance that maps ELI's to Point-of-Purchase categories allows each sampled ELI to be assigned for price collection to the outlet sample selected for the POPS category that contains it. The number of price quotes for an ELI in each outlet equals the number of times the ELI was selected for pricing in the PSU replicate during the item sampling process.

The item/outlet sample merge determines the number of price quotes assigned for collection in each sample outlet. In the outlet sampling process, if the expenditures reported for the outlet are large, an outlet may be selected more than once for a given POPS category. An outlet may also be selected for more than one POPS category. If an outlet is selected multiple times for a given POPS category, the same multiple of price quotes is assigned for collection for each sample ELI matching the category. If an outlet is selected for more than one POPS category, price quotes are assigned for collection for all ELI's selected in each category.

Selection procedures within outlets. A BLS field representative visits each selected outlet. For each ELI assigned to the outlet for price collection, the field representative uses a multistage probability selection technique to select a specific item from among all the items the outlet sells that fall within the ELI definition. The field representative first identifies all of the items included in the ELI definition and offered for sale by the outlet. When there is a large number of items in the ELI, they group them by common characteristics, such as brand, style, size, or type of packaging. With the assistance of the respondent for the outlet, the field representative assigns probabilities of selection to each group.

The probabilities of selection are proportional to the sales of the items included in each group. The field representatives may use any of four alternative procedures for determining the proportion of sales. In order of preference, they are:

a. Obtaining the proportions directly from a respondent;
b. Ranking the groups by importance of sales as indicated by the respondent and then obtaining the proportions directly or using preassigned proportions;
c. Using shelf space to estimate the proportions where applicable; and
d. Using equal probability.

After assigning probabilities of selection, the field representative uses a random number table to select one group. The field representative then identifies all items included in the selected group, forms groups of those items based on the in-common characteristics, assigns probabilities to each group, and uses a random number table to select one. The field representative repeats this process through successive stages until reaching a unique item. The field representative describes the selected unique item on a checklist for the ELI. Checklists contain the descriptive characteristics necessary to identify the item among all items defined within the ELI.

These selection procedures insure that there is an objective and efficient probability sampling of CPI items other than shelter. They also allow broad definitions of ELI's so that the same tight specification need not be priced everywhere. The wide variety of specific items greatly reduces the within-item component of variance, reduces the correlation of price movement between areas, and allows a substantial reduction in the number of quotes required to achieve a given variance. Another important benefit from the broader ELI's is a significantly higher likelihood of finding a priceable item within the definition of the ELI within the sample outlet.

This selection process is completed at the visit to the outlet to obtain the price for the selected item. Overtime, subsequent personal visits or telephone calls are made either monthly or bimonthly, to ascertain that the item is still sold and to obtain its current price.

Item and outlet samples: Shelter

The CPI housing unit sample is the source of information on price change for the two principal shelter indexes—the residential rent index and the owners' equivalent rent index. The shelter indexes account for approximately 25 percent of the total CPI weight. The housing unit sample is a stratified, systematic, multistage, cluster sample that was designed to consist of approximately 40,000 rental units and 20,000 owner units. BLS selected housing units constructed before 1980 with data developed from the 1980 Census of Population and Housing. For housing constructed since 1980, the Bureau of the Census supplies an annual sample of new units from building permits data.

Stratification. BLS used two variables, average rent level and tenure (percent renter occupied), to select the stratifying area clusters called segments; these variables correlate with rent change. Using them for sampling stratification served to make the sample sizes within clusters more con-

sistent, uniform, and homogeneous.

Stratification accomplished two goals. First, stratifying by variables associated with rent change insured sample coverage for important characteristics that correlate with rent. Second, stratification by percent renter occupied produced the clustering and the consistent sample sizes of renter and owner housing units within clusters. It is this geographic clustering that permitted the assignment or "matching" of renter-occupied units in the sample to the owner units in the sample. Matching is the mechanism that provides the basis for measuring price change of owner housing that is used in the owners' equivalent rent index.

To meet the stratification goals, BLS stratified at the lowest published Census areas within the 94 (88) CPI PSU's. In many cases, this was the block group or block level. In CPI areas where the Census Bureau provided data only by enumeration district (ED), BLS stratified at the ED level. Before stratifying areas where both block group and block data was available, BLS defined partial block groups (PBG's). Individual blocks within a block group were established as independent PBG's when they had a high percentage of renters and a large enough number of housing units to stand as an individual cluster. The balance of each original block group was also designated in total as a PBG. Less than 5 percent of the block groups were affected by this process. The use of this process, however, significantly reduced the cases where large numbers of renters were selected in a single building. The resulting PBG's were generally far more homogeneous in terms of percent renter occupied and structure type than the original block groups.

Strata boundaries were defined, and the PBG's and ED's were sorted into the strata defined on the basis of the two variables, average rent level and tenure (percent renter occupied). Eighteen strata were defined for each PSU, using three rent ranges and six tenure ranges. An important enhancement from previous CPI housing samples was that strata boundaries were defined differently for each PSU to insure that each stratum contained roughly the same number of housing units and allowed for between PSU differences in rent levels and housing characteristics. BLS sampled PBG/ED's within each stratum, thereby insuring that the survey included housing clusters of all rent and tenure levels. Stratifying by tenure also permitted BLS to vary the sample rate for owners and renters in each cluster to obtain consistent sample sizes by tenure within the clusters.

Sample allocation to PSU's and strata. BLS allocated the sample to minimize a value Z, which is proportional to the sum of the variance of the rent and owners' equivalent rent indexes. This value is expressed as:

$$Z = \sum_{i=1}^{s} \left[O_i^2 \left\langle \left(\frac{\sigma_w^2}{o_i} \right) + \left(\frac{L}{r_i} \right) \right\rangle + R_i^2 \left(\frac{L}{r_i} \right) \right]$$

where:

- s = number of strata,
- 0_i = number of owner units in the i^{th} strata,
- R_i = number of renter units in the i^{th} strata,
- L = the total unit variance,
- = within cluster variance,
- o_i = number of owners allocated to i^{th} strata,
- r_i = the number of renters allocated to i^{th} strata.

BLS determined the strata sample sizes, 0_i and r_i, subject to the sample size constraints, by finding the values for 0_i and r_i that minimize Z.[19] This produced an optimal sample for the given resource constraints for the two indexes.

BLS determined an initial allocation simultaneously across all strata and PSU's based on a criterion requiring a minimum sample size for each published index. If a publication PSU was not allocated the minimum sample, it was assigned a designated number of units large enough to meet publication standards. This minimum sample was allocated applying the above formula for Z among the strata within the PSU. A single process reallocated the remaining sample to the remaining PSU's.

The budget for the CPI Housing survey dictated a sample of 10,000 clusters (called segments) and 100,000 pricings per year for the pre-1980 sample portion of the survey. BLS added 900 more segments to compensate for an expected 9-percent sample loss that results from differences in the Census Bureau and CPI definitions of housing units. In contrast to the Census definition, the CPI excludes public housing, institutional housing, and military housing.

Sample selection within strata. BLS selected sample clusters in each stratum using a systematic probability-proportional-to-size (PPS) sample selection method. Each partial block group within a stratum was assigned a measure of size according to the total number of housing units, with controls on the maximum and minimum percent renter occupied. BLS sorted the PBG's geographically and, using the measures of size, allocated the sample of PBG's systematically.

Next, BLS partitioned each selected PBG into a number of clusters, depending on the PBG's size, and selected one at random. When a single Census block contained more than one cluster and one of these was a selected cluster, BLS field representatives defined the individual cluster following strict procedures derived from the sampling plan. For example, suppose that the sample design determined that a segment began 10.1 percent of the way into a block and ended at 23.8 percent into the block. If the block was not too big, the BLS representative would enumerate the entire block and then define the segment using these percentages. In large blocks,

[19] For a full derivation of Z, see W.F. Lane and J.P. Sommers, "Improved Measures of shelter Costs," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 1984.

the field representative prescreened the blocks and sent the information to the BLS national office which determined the segment.

A field screening determined the final selection of housing units in the sample. In the first step of this process, called listing, BLS field representatives enumerated in order on listing forms every housing unit or potential housing unit they saw in each segment. The national office prepared unique selection sheets that indicated the sequence number of each unit to be screened and its "desired tenure" (whether the unit needed to be owner or renter occupied to pass screening and remain in the survey). Using these selection sheets, the field representatives identified on the listing forms the units to be screened and their desired tenure.

The selection sheets for each segment took into account each segment's proportion that was renter occupied according to the 1980 census to determine how many units to screen, and how many to require of each tenure to yield an approximately optimal sample of renter and owner units for the CPI Housing survey. The two within-segment sampling rates determined how many renters and owners should be in the final sample. These rates were based on the final desired probability of selection for each tenure, the number of renters and owners in each stratum, and the number of selected segments and total segments in each stratum.

Thus, before the field agents contacted survey respondents for the first time, the units in each segment had already been assigned to one of four cases:

(1) Screen but initiate only if the unit is renter occupied;
(2) Screen but initiate only if the unit is owner occupied;
(3) Screen and initiate if either owner or renter occupied; or
(4) Do not contact for screening or initiation.

During screening, the field representatives contacted an eligible respondent for each housing unit to be screened and determined that the unit met the tenure criteria for the survey as well as other criteria such as being a year-round housing unit, built before 1980, and that it was someone's primary residence. Units that passed screening were immediately initiated into the sample.

Initiation is the process of collecting a first-time interview. Units that did not pass screening were not initiated; however, they may be recontacted at a later date to augment or to rotate the sample.

As planned, only about one-fourth of the units interviewed for screening met the tenure and other eligibility criteria required to pass screening and be initiated into the CPI Housing survey. Because renters were allocated according to total units and each cluster was allocated an equal number of renters, most of the units contacted in areas known from the Census Bureau to be mostly owner occupied failed screening. However, this process located the sparse renters in these areas for the survey and added ''extra" renter units to the rent sample in owner areas. Although they represent few renters in the renter universe and, consequently, have very low weight in the rent index, they serve as the main source of rental units to match with owner units. It is their movement that primarily drives the owners' equivalent rent index.

Sample augmentation and rotation. In 1989, BLS augmented the renter sample in those strata containing a balanced mix of owner and renter units. Additional renter units in segments within these strata were selected, screened, and initiated, and the sampling weights for these segments were appropriately adjusted. The augmentation resulted in a 5-percent increase in the overall number of renters in the CPI housing survey and specifically added units that support both the rent and owner's equivalent rent indexes.

In 1992, BLS began rotating/rescreening segments in heavily owner-occupied strata, primarily to support the owner's equivalent rent index. Sampled rental units in heavily owner-occupied areas have demonstrated a substantial propensity to change tenure from renter to owner during the past 5 years. Through rotation, BLS will regain renters in these areas by rescreening these segments to locate units that were owner-occupied during the original screening process but have since become rentals. This rotation will also result in a new sample of owner units in these segments.

Estimation of Price Change

Commodities and services other than shelter

At the end of each pricing period, the estimate of the one-period (t-1 to t) price change (price relative) is computed for each item-area stratum. Only price quotes obtained in both the current and previous pricing periods for the same or comparable items are used in the estimate. Where appropriate, price items are converted to a price per unit of measure before they are used in the estimation of price change. The same quote weights are used both for the current- and previous-period price quotes. The estimate of the one-period price change for the h^{th} index area for the z^{th} item stratum for a given market basket is computed as:

$$R_{hzt,1} = \frac{R_{hzt,a}}{R_{hzt-1,a}} = \frac{\sum_{i \varepsilon z} W_{hi} P_{hit} / P_{hia}}{\sum_{i \varepsilon z} W_{hi} P_{hit-1} / P_{hia}}$$

where:

P_{hit} is the price of the i^{th} quote in the current pricing period, t, for item stratum z in index area h;

P_{hit-1} is the price of the i^{th} quote in the previous pricing period, t-1, for item stratum z in index area h;

P_{hia} is the estimated price for the i^{th} quote for item

stratum z in the time period, a, of the POPS in index area h;

W_{hi} is the quote weight for the i^{th} quote for item stratum z in index area h. It is computed as:

$$W_{hi} = \alpha Efg / MB$$

where:

- α is the percent of sales of the ELI to the total sales of the POPS category in the outlet;
- E is an estimate of the total daily expenditure for the POPS category, for the index area replicate and the CPI-U population;
- f is a duplication factor to reflect any special sub sampling of outlets or quotes;
- g is a geographic factor to reflect the difference in coverage for the index area for the pre-1987 area definitions to the 1987 area definitions;
- M is the number of usable quotes for the ELI/PSU replicate within the item stratum; and
- B is the proportion of the expenditures for the selected ELI is of the total expenditures for the item stratum in the region (the probability of selection for the ELI).

NOTE: The geographic factor is 1.000 for all samples selected using the 1987 area definitions. See the section on outlet sampling procedures for an explanation of sample rotation.

Item substitution, quality adjustments, and imputation

One of the more difficult conceptual problems faced in compiling a price index is the accurate measurement and treatment of quality change due to changing product specifications and consumption patterns. The concept of the CPI requires that BLS measures through time, the cost of purchasing a fixed, constant-quality market basket of goods and services. In reality, products frequently disappear, products are replaced with new versions, and new products emerge. BLS uses several methods to adjust for quality change and to account for the change in item specifications. These methods may be categorized as 1) directly comparable, 2) direct quality adjustment, and 3) imputation. In all cases, it is necessary to estimate a new base-period price in order to use the new item specification in future periods. The base-period price (called P_{hia} in the above formula) is an estimate of the item's price at the time it was reported on the POPS.

Directly comparable. If the new and old item specifications are considered directly comparable, i.e., the characteristics of the new item are essentially the same as the old item's characteristics, the base-period price for the new specification is set equal to the base-period price for the old specification, and the price comparison between the items is used in the index. It is assumed that no quality difference has occurred.

Direct quality adjustment. This is the most explicit measure for dealing with specification changes. Direct quality adjustments are frequently made for the automobile and apparel components of the CPI.

The most frequently cited example of direct quality adjustment is the annual model changeover for new cars and trucks. Each year, price adjustments are made to account for the quality differences between the old and the new models. In some cases, the adjustments are based on the previous model's retail price for optional equipment. In other cases, the quality adjustments must be derived from production cost data supplied by the manufacturers. These data are adjusted by estimated manufacturer and retailer markup rates to derive retail values for the quality changes.

Adjustments for quality change in the CPI new car index include structural and engineering changes that affect safety, environment, reliability, performance, durability, economy, carrying capacity, maneuverability, comfort, and convenience. Although antipollution equipment on automobiles does not directly increase the quality of the automobile for the buyer, these devices do improve the quality for consumers in general. Consequently, quality adjustments are made for pollution controls to automobiles on the assumption that, by legislative definition, the cost of installing antipollution devices was no more than the value derived from them.

Quality adjustments of new cars exclude changes in style or appearance, such as chrome trim, unless these features have been offered as options and purchased by customers. Also, new technology sometimes results in better quality at the same or reduced cost. Usually no satisfactory value can be developed for such a change. In such cases, it is ignored, and prices are compared directly.

In addition to quality adjustments for physical changes to cars and trucks, adjustments are made for changes in the warranty coverage provided by auto manufacturers when sufficient data are available to derive estimates of their values.

The marketing of apparel items have historically caused a number of problems in the maintenance of a constant quality market basket of apparel commodities in the CPI. Many apparel items are seasonal and subject to frequent style changes. In addition, heated competition in the marketing of apparel commodities has led to increasingly rapid turnover of styles available in retail outlets. Until recently, these factors have meant that, when new styles replace old ones, many substitutions were deemed not comparable. Marketing practices for apparel generally entail introducing such goods at high regular prices and marking them down to lower "sale" prices throughout their shelf life. Therefore, the inability to compare prices at the time the new goods

are introduced (when price increases are normally passed along to the consumer) precludes capturing any pure price change which accompanies the style changes. As a result, during the early to mid 1980's, apparel indexes understated price change. Research undertaken to correct this problem delineate the critical determined apparel quality characteristics to hold constant in the substitution process. When an outlet discontinues an apparel item, the field representative follows the CPI substitution procedures to find the closest substitute it offers for sale. The procedure insures that as many as possible of the critical quality characteristics of the substitute are the same as those of the discontinued item. For those that are not the same, CPI apparel experts estimate the value of the difference and directly adjust the price to account for them.

Hedonic regression modeling is the technique which was used to determine the importance of the various quality characteristics which add value to a particular good. In this approach, an item can be viewed as a collection of characteristics which, taken together, provide satisfaction or value to the consumer. In other words, a women's suit can be considered an aggregation of its components, such as a jacket and skirt or pants, each of which contributes to the value of the suit in the eyes of the consumer. In addition, certain aspects of the suit, such as its fiber content, its construction, etc., will add or detract value from the consumer's standpoint. Hedonic regression modeling is a tool which allows the researcher to estimate which factors are the most important price-determining characteristics of these goods. In the CPI, this research resulted in better data collection documents and procedures for pricing apparel commodities. By noting the most important quality characteristics on the data collection document, the field agents who collect data for the CPI can try to hold constant these important characteristics even when other details change. This improvement in the collection documents resulted in a remarkable increase in the number of comparable substitutions chosen for apparel commodities.

As additional research led to further improvements in the modeling process, steps were taken to further reduce the number of substitute items deemed not comparable to the old items. Tests on the data which are collected for apparel goods determined that direct quality adjustments could be made using the estimates of the values of the different components of the apparel items. If, for example, a two-piece men's suit had been priced but was no longer available, all the two-piece suits having been replaced by the new three-piece style, the value of the vest can be added to the price of the old two-piece suit and the quality-adjusted price of the old suit compared directly with the price of the new three-piece suit. Alternatively, if fiber percentages vary between two items, quality adjustment can account for the quality difference to permit comparison of the prices of the two goods. Other differences which have been shown to be statistically significant can also be factored out to permit constant quality price comparisons of apparel items. This critical change allows another incremental step forward in decreasing the number of substitutions which can not be compared due to differences in quality characteristics. While this change has its greatest impact in apparel, since apparel has been plagued with low comparability rates, this same research is underway in all other areas of the CPI so that marginal improvements in comparability in other areas can be expected in the next few years.

The hedonic approach to quality adjustment used for apparel is in contrast to the approach used for new cars, which is based on manufacturers' costs. While those estimates of the value of the quality change are based on the manufacturers' cost differentials adjusted for retail mark-ups, the quality adjustments in apparel are based on the average consumers' valuation of the change as well as on the manufacturers' costs to produce the goods. This valuation is what the hedonic modeling technique—based on the prices paid by consumers for the goods for which the CPI collects prices, implicitly estimates. As this research continues, BLS expects to make further improvements in the quality adjustment process.

In general, if the new item specification is similar to the previous one but has changed one or more of its component parts, a quality adjustment can establish comparability between their prices. A synthetic previous-period price for the new item ($P^*_{i,t-1}$) is calculated as follows:

$$P^*i,t\text{-}1 = Pi,t\text{-}1 + QA$$

where:

$P_{i,t-1}$ is the actual previous-period price of the previous item and QA is the dollar value of the quality change, which may be either positive or negative.

After the above calculation is made, the base-period price for the new item ($P_{*i,a}$) is computed as:

$$P^*_{i,a} = \frac{P_{i,a} P^*_{i,t-1}}{P_{i,t-1}}$$

where:

$P_{i,a}$ is the base-period price for the previous item.

For new cars and trucks, where the quality adjustment is expressed in current dollars, $P^*_{i,t-1}$ is calculated as follows:

$$P^*i,t\text{-}1 = Pi,t\text{-}1\ [Pi,t / (Pi,t - QA)]$$

Imputation. For some item specification changes, however, BLS has not yet determined how to make quality adjustments. Substitute items that can neither be directly compared nor quality-adjusted are called noncomparable. For noncomparable substitutions, an estimate of constant-qual-

ity price change is made by imputation.

There are two types of imputation used for commodities and services. The first type of imputation is prevalent for noncomparable substitutions that occur for food and service items. In this type of imputation, the rate of price change between the old item and the noncomparable new item is assumed to be the same as the average price change of all similar items used in CPI calculations that month for the same geographic area.

To execute this imputation or "link," two estimates of price change for the item-stratum are required: ($R_{zt-1,a)}$ the long-term change from the CPOPS period a, to the previous pricing period t-1 and ($R_{zt,t-1}$) the 1-month change from the previous period to the current month. A new base-period price ($P^*_{zi,a}$) for item i, in item stratum z is computed as follows:

The value of $R_{zt-1,a}$ is estimated at either the item level by using the ratio of the previous-period price to the base

$$P^*_{zi,a} = \frac{P_{zi,t}}{R_{zt-1,a} * R_{zt,t-1}}$$

price, or at the item stratum level by using the ratio of indexes for the item-area stratum where:

$$R_{zt-1,a} = \frac{I_{z,t-1}}{I_{z,a}}$$

$I_{hz,t-1}$ is the index of the previous pricing period for the item-area stratum hz,

$I_{hz,a}$ is the base-period index for the CPOPS period (or other frame reference), and

$R_{hzt,t-1}$ is the one-period price change relative for the item-strata stratum.

The quality difference between the items in this case is assumed to be the difference between the price of the new variety and the imputed price for the old variety.

For many other items, however, price change is closely associated with the annual or periodic introduction of new lines or models. For example, price changes for new cars often accompany the introduction of new models. Price changes for cars within the same model year, then, are not the best estimate of price change for noncomparable substitutions that occur between model years. Rather, the average price change that occurs between model year changeovers used in CPI calculations provides a superior approximation of price change for noncomparable new car models. Therefore, since October 1989, price changes for noncomparable new model vehicles have been imputed using only constant-quality price change of comparable model changeovers. Price changes recorded within the same model year are excluded from the estimate. This type of imputation using price changes for substitutions of comparable quality to estimate price change for noncomparable substitutions will become increasingly prevalent for nonfood commodities.

In order to impute noncomparable substitutions by the price change for comparable substitutions, a current-period "overlap" price ($P^*_{i,t}$) is estimated for the old specification of the item as follows:

$$P^*i,t = Pi,t-1\ Rct,\ t-1$$

where $R_{ct,t-1}$ is the current one-period relative for all comparable and quality-adjusted substitutions in the same item-area stratum.

When a price is obtained for the old ($P_{i,t}$) item specification, and imputed for the new ($P^*_{i,t}$) specification in the same period (overlap pricing), the estimation of the base period price is based on the same-period price relationship of the two specifications. The new base-period price ($P^*_{i,a}$) is estimated as follows:

$$P^*\ i,a = P^*\ i,t\ (Pi,a\ /\ Pi,t)$$

For the current month, the price comparison used in the index is $P_{i,t-1}$ to $P^*_{i,t}$ At the next pricing, the comparison will be made on the new item. The quality difference is assumed to be the difference in the prices of the old and new varieties in the current month.

Medical care. Another area in which quality adjustment presents particular difficulties is medical care. Not all factors affecting the quality of medical care services can be accounted for in the description of the item being priced. Quite often the respondent does not have knowledge of many price-determining quality factors. For example, hospital room modifications, changes in the nurse-to-patient ratio, or the availability of new equipment are all likely to contribute to determining the price level of the room service priced. Such changes are normally reflected as price movement because BLS either is not aware of the changes or has no method available to deal with the change. Improved technologies and procedures can lead to quality changes that cannot necessarily be measured by BLS. For instance, new advances in the development of porous materials in the manufacturing of prosthetic implants, such as in hip replacement surgery, allow the bone to grow around the prosthesis. This is not the case with the nonporous materials that have been commonly used in hip replacement prosthetic implants. Many doctors view this porous implant as an improvement in the results of hip replacement surgery. In pricing total hip replacement surgery, the quality impact of shifting from a nonporous to a porous implant would not be factored out of the index, as BLS has no methodology to account for the value of the quality difference.

There are, however, certain areas in medical care where the quality difference can be measured and adjustments made

for changes in the quality of priced services. For example, the CPI might be pricing a limited visit to a physician's office for treatment of a sore throat, and the physician changed the fee schedule so that a throat culture was always performed and the price included in the cost of the visit. The addition of the throat culture would not be reflected as a price increase, because it was not in the described service and would be considered a substitution. If the physician identifies the cost of the office visit for the current month, a price change in the cost of the office visit can be reflected from the previous month. For subsequent pricing comparisons, the new service would include the office visit and throat culture. If a hospital introduces a separate admitting charge that previously was included in the room rate, BLS prorates the admission charge to a per-day basis using an appropriate hospital-provided length-of-stay measure. The prorated admission charge is then added to the room rate priced to reflect the price movement in the index.

Other price adjustments

Bonus merchandise adjustments. Sometimes products are offered with free merchandise included with the purchase of the original item. Such ''bonus" items may provide additional satisfaction to consumers, and BLS will, therefore, make adjustments to the purchase price to take into consideration the value of the bonus merchandise. The adjustment made depends on the type of merchandise offered and the perceived value of the bonus to the consumer.

If the bonus merchandise consists of more of the same item, the adjustment is reflected in the price of the item. For example, if a manufacturer offers two ounces of toothpaste ''free with the purchase of the regular six-ounce tube," then the item's price will be adjusted to reflect a decrease in the per-ounce price. When the bonus is removed, the price per ounce will return to its prior level, and a price increase will be recorded. In this instance, the value to the consumer is assumed to be one-third greater.

If the bonus merchandise consists of an item that has some significant value to the consumer, and the item is of a different genre, an adjustment will be made to account for the value of the free item when feasible to do so.

When bonus merchandise adjustments occur, base prices are not adjusted since there is no difference in the quantity or quality of the original item being priced.

Utility refunds. Sometimes public utility commissions require that utilities such as telephone, natural gas, or electricity companies make rebates to their customers. These rebates may arise from a number of different causes. For example, a utility may be permitted to use a new rate schedule temporarily until a final determination is made. If the final rates set by the commission are lower than the temporary ones, the difference must be refunded for consumption during the period. The utility bills priced for the CPI will reflect the full amount of these refunds in the month(s) they are credited to the customers.

Manufacturers' rebates. When product manufacturers offer cash rebates to consumers for purchases of items priced in the CPI, these rebates will be reflected in the index as price reductions. The amount of the rebate adjustment usually depends on the percentage of purchasers who take advantage of the rebate offer. For example, when auto manufacturers offer a $500 rebate on the purchase of a new car or reduced-rate financing, the price of each car eligible for the rebate is reduced by the proportion of customers who opt for the rebate. If 70 percent of customers choose rebates for a particular model, then the price of each quote for that model in the CPI will be reduced by $350, and the index will reflect the price decline. The reduced interest rates chosen by the remaining customers will be reflected in the auto financing component of the CPI. For mail-in rebate offers, an attempt is made to determine the proportion of customers who take advantage of the rebate, and the reported prices are adjusted accordingly.

Cents-off coupons. Generally, no adjustments are made for coupons presented by customers as price reductions at the time of payment. One exception is when the coupons are attached to the product for immediate redemption at the point of purchase. BLS field representatives are instructed in this latter situation to deduct the amount of the coupon from the price of the item.

Seasonal items

Seasonal items are those commodities and services that are not available year round but are available only at certain times of the year. Heavyweight coats, tents, and fresh peaches are examples of seasonal items. Special procedures are employed when selecting and pricing these types of items to ensure they are appropriately represented in the sample and price changes are correctly included in the calculation of the CPI. In particular, the procedures prevent substituting away from a seasonal item when it is out of season.

Although seasonal items can exist in any ELI, some ELI's include an especially large percentage of items that are seasonal and, consequently, receive special treatment. These seasonal ELI's include most apparel items, fresh fruit, and sports and recreational equipment. The designation of an ELI as seasonal or non-seasonal was made at the regional level, using the four geographic Census regions in the CPI design. It is not uncommon for some items that are seasonal in the Northeast region, for example, to be sold year round in the South.

After the samples for these seasonal ELI's are selected following the normal sample selection procedure, the number of quotes is doubled to ensure that, despite the seasonal disap-

pearance of a large number of quotes, a large enough number of in-season quotes will remain to calculate the index.

The quotes in these ELI's are paired; that is, for each original quote that is selected, a second quote in the same ELI and outlet is initiated and priced 6 months later. In the fresh fruit ELI's, one quote of each pair is designated January-June, and the other quote is designated July—December. In all other seasonal ELI's, one quote of each pair is designated fall/winter, and one quote is designated spring/summer. The fall/winter and spring/summer designations are used for the nonfood quotes because these are the distinctions that are most commonly used by the retailing industry to categorize seasonal merchandise. These seasonal designations are used to help establish the specific items eligible for each quote so that year-round items and items from each season are initiated in their proper proportions.

For every specific commodity and service priced in the CPI, including year-round items, BLS field representatives collect—at the time they initiate the item and every time they must find a substitute for it—its in-season months. These data become a part of the item description and are updated if there is a change. Field representatives attempt to price every item in each period it is designated for collection, even during those months when the item may be out of its indicated season. If the item is available, the price is collected and used in the calculation of the CPI. If the item is unavailable because it is out of season, no substitute item is selected nor is any further action taken, and that item is not used in the calculation of the CPI that month. Its price is imputed using standard imputation procedures.

When an item becomes permanently unavailable, the standard procedure is to substitute the most similar item. In the case of a year-round item not in a seasonal ELI, this process takes place as soon as the item is permanently unavailable. For items in seasonal ELI's and seasonal items in ELI's that are not designated seasonal, however, the period during which a substitution can take place is restricted to those months when a full selection of appropriate seasonal merchandise is available.

These special initiation, pricing, and substitution procedures are intended to ensure that an adequate sample of items is available every month, and the correct balance of seasonal and year-round items is maintained. As a result, the estimates of price movement for the ELI's that include seasonal items are hoped to reflect price changes for the universe of items included in those ELI's correctly.

Special estimation procedures

A number of special estimation procedures are used in compiling price information for selected categories of items in the CPI.

New vehicles. Prices for new cars and trucks selected for inclusion in the CPI pose a special problem since the manufacturer's suggested retail (sticker) price does not represent the transaction price for most new vehicles. Most automotive dealers offer customers concessions on the sticker price, or, at those times when models are in high demand, the dealers will charge an additional markup beyond the sticker price. When pricing new vehicles, BLS field representatives obtain separately the base price and the prices for options, dealer preparation, transportation, etc. In addition, they obtain from the dealer the average concession and/or markup during the preceding 30 days. This enables BLS to estimate the true transaction price for each vehicle after concessions/markups.

Used cars. The only expenditures on used cars included in the CPI market basket are those for previously owned cars consumers purchased from the business or government sectors and the profit of dealers on the sale of used cars. (See special expenditure weight issues above.) The used car sample was selected from types of cars purchased for use by businesses and governments. The sample consists of 2-through 6-year-old models. Average wholesale prices of clean cars sold at used car auctions are published by the National Automotive Dealers Association. The average of these prices is adjusted for depreciation using the difference in prices between model years for the same model car. The prices used in the index are a 3-month moving average of the average wholesale prices adjusted for depreciation.

Natural gas. The energy value of natural gas varies according to the quality of the gas supplied. BLS attempts to price a constant amount of energy consumption for natural gas. When natural gas is sold by volume—e.g., cubic feet—the amount of gas needed to produce a constant amount of energy will vary depending on the heating value of the gas. To ensure that a constant amount of energy is being priced, the amount of gas consumed is adjusted each month based on the current heating value. Thus, through time, a constant amount of energy is priced. The current adjusted consumption is calculated as follows:

Current adjusted consumption = original consumption x (original heat value/current heat value).

Health insurance. Health insurance is not directly priced in the CPI. The price change is imputed from the price movement of the various services that are covered by health insurance and from the change in the ratio of retained earnings to benefits paid by type of health insurance carrier - Blue Cross/Blue Shield or other. (For additional detail, see special expenditure weight procedures.) Thus, the price change for health insurance, by type of carrier, is estimated each month by the product of two relatives—one for the change in the various assigned medical care items (e.g., physician services, hospital rooms, etc.) and the other for the change in the retained earnings ratio of the carrier.

Automobile finance charges. The price used in the CPI for automobile finance charges represents the amount paid for financing a loan with fixed characteristics such as down payment percentage, term of the loan, type and model of car, etc. The price change is affected by two items—the interest rate on the loan and the changes in the amount financed due to price movement for new cars. The automobile financing charges index is estimated each month by the product of two relatives, one for changes in interest rates charged on new-car loans and the other for changes in new-car prices.

Quantity discounts. Many items in the CPI are sold both individually and in quantity. When consumers are able to purchase an amount greater than a single unit at a discounted price, the first multiple unit price is reported for use in the CPI. For example, if the 12-ounce can of corn being priced can be purchased at 25 cents for a single can, three cans for 69 cents, or five cans for $1, the price used in the CPI will be the per-ounce price of the three cans.

Unit-priced fruits and vegetables. When pricing fresh fruits and vegetables that are sold on a unit basis, two of the items are weighed to determine an average weight for the item. This helps to reduce the variability in the size that occurs among individual, loose-produce items and is not overly burdensome for the data collection process. For example, if the item being priced is Red Delicious apples and the price is 50 cents each, the BLS field staff will report the price of one apple and the combined weight of two Red Delicious apples taken from the produce rack. In computing the price per ounce, the weight of the two apples is divided by two and the price of an apple is divided by this average weight.

Bottle deposits. BLS collects information on bottle deposits for a variety of nonalcoholic and alcoholic beverages in order to calculate the influence of changes in bottle legislation on price change. Consumers who purchase throw-away containers are considered to be purchasing both the product itself and the convenience of throwing the container away. When a local jurisdiction enacts bottle legislation and no longer allows stores to sell throwaway containers, those consumers who were previously purchasing throwaway containers may experience a change in the price of the convenience. The price of the same size bottle of product plus its deposit establishes an upper bound for the price change since the consumer could retain the former convenience by now purchasing returnables and simply throwing them away. In similar fashion, information about deposits and the status of bottle legislation can be used to estimate price change when a bottle bill is repealed. Changes due to bottle bills are shown in the month the legislation is effective.

Sales taxes. The CPI includes all applicable taxes paid by consumers for services and products purchased. A number of prices for services and products used to calculate the CPI are collected with taxes included because this is the manner in which they are sold. Other prices are collected excluding applicable taxes with those taxes subsequently added in the Washington office. The tax rates for these items are determined from secondary sources based on the State and local tax structure governing the sale of the service or product at the point of purchase. BLS currently is studying the effects of taxes on the monthly movement of the CPI.

Shelter: Residential rent and owners' equivalent rent

The residential rent and owners' equivalent rent indexes measure the change in the cost of shelter for renters and owners, respectively. Price change data for these two indexes come from the CPI Housing survey. Each month, BLS field representatives gather information from renter units on the rent for the current month the previous month and on what services are provided. From owners' units, they obtain an estimated or implicit rent and from all units they collect information on characteristics of the sample housing units and respondents.

Residential rent. The rent estimates used in the CPI are "contract rents:" they are the payment for all services the landlord provides in exchange for the rent. For example, if the landlord provides electricity, it is part of the contract rent. The CPI item expenditure weights also include the full contract rent payment. The CPI rents are calculated as the amounts the tenants pay their landlords plus any rent reductions tenants receive for performing services for the landlord (sometimes called "rent as pay") plus any subsidy payment paid to the landlord. If the rent is lower than prevailing market rents because the renter is related to the landlord, the unit is not used in the calculation. Reductions for any other reasons are not considered part of the rent.

BLS makes two preliminary estimates of the rent index for each CPI area: Its current month's index is estimated, first, by applying estimates of 1-month change to its next index for the previous month and, second, by applying estimates of 6-months rent change to its index for 6 months earlier. The estimate of the 1-month rent change is the sum of the current month's rents—weighted and adjusted for 1 month of aging—divided by the previous month's sum of weighted rents. The estimate of the 6-month rent change is the sum of the current month's rents—weighted and adjusted for 6 months of aging—divided by the sum of weighted rents for the previous 6 months. The final rent index for the current month is a weighted average of the previous month's rent index moved forward by the estimate of 1-month rent change and the rent index from 6 months earlier advanced by the estimate of 6-month change.

To put this in the form of an equation, let S_1 be the set of

rental units in the CPI Housing survey in a CPI area with valid comparable rents in both time t and in time t-l; and let S_6 be the set of units interviewed in time t with valid comparable rent values in both time t and time t-6. Vacant units that were previously renter occupied are also included in S1 and S_6 and have current (t) and previous (t-l) month's rents assigned using a vacancy imputation process. Let the rent for rental unit i in time t be rit, and let ai be the factor, which is discussed below, that adjusts for the estimated small loss in quality due to the aging that the unit experienced between t-l and t. The 1-month and 6-month estimates of rent change, $R_{t,t-1}$ and $R_{t,t-6,}$ are calculated by:

$$R_{t,t-1} = \frac{\sum_{i \varepsilon S_1} \left(r_{i,t} + a_i r_{i,t-1}\right) W_{i1}}{\sum_{i \varepsilon S_1} r_{i,t-1} W_{i1}} \qquad R_{t,t-6} = \frac{\sum_{i \varepsilon S_6} \left(r_{i,t} + 6 a_i r_{i,t-6}\right) W_i}{\sum_{i \varepsilon S_6} r_{i,t-6} W_{i6}}$$

where:

W_{i1} and W_{i6} are the sampling weights, which are the verse of the renter units' probability of selection adjusted for nonresponse.

Using $R_{t,t-1}$ and $R_{t,t-6}$ and the indexes for the previous month I_{t-1}, and for the 6 months previous, I_{t-6}, BLS computes two preliminary estimates ($I_{t-1}R_{t,t-1}$ and $I_{t6}R_{t,t-6}$) of the current month's rent index, It, for each market basket. The final rent index for month t for each market basket is the weighted average of the two preliminary estimates:

$$It = A(It\text{-}1\ Rt,t\text{-}1) + (1\text{-}A)\ (It\text{-}6\ Rt,t\text{-}6)$$

where:

A=0.65, the value that simulation studies determined to minimize the mean squared error of the estimate:[20]

Vacancy imputation. Vacant units which were previously renter occupied are used in the calculation of $R_{t,t-1}$ and $R_{t,t-6}$. The vacancy imputation process incorporates several assumptions about the unobserved rents of vacant units. It is assumed that rents tend to change at a different rate for units that become vacant (and are, therefore, in the process of changing tenants) than for other units. The vacancy imputation model assumes that, after an initial lease period, expected rents change at a steady rate until the old tenant moves out of the unit. When there is a change in occupant or a unit becomes vacant, its rent is assumed to "jump" at some rate, referred to as the "jump rate." In markets with generally rising rents, this jump rate is usually greater than the average rate of change for occupied units. BLS estimates the jump rate based on non-vacant sample units in the PSU which have had a change in tenant between t and t-6. Non-vacant units without a tenant change are used to calculate the average continuous rate of change. These values are used to impute rents for vacant units for periods t and t-1 from their rent in t-6.[21] In general, the imputed rents, ri,t and ri,t-1, of the i^{th} vacant rental unit in t and t-1 are:

$$ri,t\text{-}1 = ri,t\text{-}6\ C5 \text{ and } ri,t = ri,t\text{-}1J$$

where is the jump rate for the PSU calculated, and C is the steady rate of change.

The imputation of vacant rents ensures that the unobserved rent change that occurs when a unit becomes vacant will be reflected in the final rent index. The 6-month rent-change estimates will capture these changes once the units become occupied, however, they will be missed in 1-month rent-change estimates without vacancy imputation. Because the final rent index is calculated using both 1- and 6-month change estimates, omission or misstatement of rent estimates for vacant units would lead to BLS missing part of rent changes in the CPI.

Aging adjustment. The aging adjustment accounts for the small loss in quality as housing units age (or depreciate) between interviews. The aging adjustment factors, a_i, can be thought of as 1/(1-d) where d is the monthly rate of physical depreciation. BLS computes factors for each housing unit with regression-based formulas. The formulas account for the age of the unit and a number of structural characteristics.[22] The aging adjustment procedure was introduced into the CPI in 1988.

Quality adjustment. Quality adjustments made to the cost of rental housing are used in the rent and owners' equivalent rent indexes. In addition to collecting the rent charged, BLS also collects a description of major services and facilities provided by the landlord. If the services and facilities differ between two collection periods when rents are compared, the rent for the current period is adjusted to reflect the differences in services between the time periods. For instance, if the owner no longer provides a certain utility, BLS calculates an estimate of the value of that utility and adds it to the current rent in order to have an adjusted rent value. This adjusted rent is the current cost of the same set of services provided for the previous rent.

[20] For a derivation of the optimal value of A, see C.L. Kosary, J., P. Sommers, and J. M. Branscome, "Evaluation Alternatives to the Rent Estimator," *Proceedings of the Business and Economic Statistics Section,* American Statistical Association, 1984.

[21] For more information on vacancy imputation, see J.P. Sommers and J.D. Rivers, "Vacancy Imputation Metrology for Rents in the CPI," *Proceedings of the Business and Economic Statistics Section,* American Statistical Association, 1983

[22] For further information, see Lane, Walter F.; Randolph, William C., and Berenson, Stephen A; "Adjusting the CPI Shelter Index to Compensate for Effect of Depreciation,'' *Monthly Labor Review,* October 1988, pp. 34-37.

To make quality adjustments in costs of utilities, BLS uses data from the Department of Energy's Residential Energy Consumption Survey to develop formulas to estimate utility usage for various types and sizes of housing, in various climates, with different types of heating and air conditioning, hot water, and so on. Prices for utilities come from the CPI average price program. A similar, simpler formula is used to estimate water costs.

Using data calculated yearly from the aging adjustment regressions, quality adjustments for major structural changes (i.e., the number of bedrooms, bathrooms, or other rooms and central air conditioning) have been made since February 1989.[23] Previously, when such major changes occurred, BLS omitted these observations from the calculation for estimation of price change for that time period.

Owners' equivalent rent. In Part I of this chapter, the concept of using owners' equivalent rent to measure homeowner shelter costs was discussed. BLS estimates the owners' equivalent rent index[24] by estimating the owners' implicit rent, m_j for each owner unit, j, in the sample. In contrast to the contract rent concept used in the rent index, the implicit rent is a "pure rent"; that is, it excludes payments for extra services such as utilities and furniture. Once the implicit rents are estimated, the calculation of the owners' equivalent rent index essentially follows that of the rent index.

The initial value derived for time t for $m_{j,t}$, is an estimate of the rent the owner-occupied housing units in the Housing survey would bring if they were rented. The estimate is based on the answer to the question,''How much do you think you could rent this house out for monthly, not including utilities?" For owner-occupied units whose owners are unable to estimate their unit's implicit rent, BLS uses estimates provided by its field representatives based on their knowledge of the rental market in that area. If this information cannot be obtained, BLS uses an imputation procedure that assigns the implicit rent from a similar unit to any that have missing values.

To get subsequent values of implicit rent each month, BLS assigns a set of renters, Q_j, to each owner unit, j. This assignment is done on the basis of location within the PSU, structure type, and structural characteristics. BLS first tries to match owners with renters that fit for all variables. For those owners for whom a matching set of renters is not obtained at the first stage, BLS relaxes the constraints one at a time until a satisfactory set of renters is found for all the owners. In general, a single renter may be assigned to sets for estimating no more than three owner equivalents. When several renters, say nj are assigned to owner j, this counts as only $1/n_j$ toward each renter's maximum of three owners. Renters are only checked against their maximum after a round of matching, however, soft is possible for a renter to move more than three owner equivalents if the renter is matched to more than three during one round. The sample selection process, which sampled renters in owner areas at a very high rate, facilitated the matching of renters to owners. (See the section on item and outlet samples for shelter.) BLS estimates the pure rent, $P_{i,t,}$ for all the rental units in Q_j. This pure rent estimate, $P_{i,t}$, is $r_{j,t}$ minus an estimate of the value of any utilities or furniture the landlord provides.

Q_{j1} is the subset of the renters in Q_j that have valid comparable rents in both t and t-1. Q_{j6} is the subset with valid rents in both t and t-6. Vacant, previously renter-occupied housing units are eligible for Q_{j1} and $Q_{j6.}$ The implicit rent, mjt for owner j in time t is estimated from the implicit rent for t-6 and the average change in the pure rent of the units:

$$m_{j,t} = m_{j,t-6} \sum_{i \varepsilon Q_{j6}} \left(\frac{P_{i,t}}{P_{i,t-6}} + 6a_i \right) \Big/ n_{j6}$$

where $P_{i,t}$ is the ith rental unit's pure rent, ai is the aging adjustment factor, and nj6 is the number of rental units in Q_{j6}.

The 1-month previous implicit rent is the current month's implicit rent moved back 1 month with the pure rents in Q_{j6}:

$$m_{j,t-1} = m_{jt} \sum_{i \varepsilon Q_{j1}} \left(\frac{P_{i,t-1}}{P_{i,t} + a_i P_{i,t-1}} \right) \Big/ n_{j1}$$

Once BLS obtains estimates of current, 1-month-ago, and 6-months-ago implicit rents for all owners, it proceeds to estimate the owners' equivalent rent index for the current month. The process is similar to that used for the rent index. There is no problem here with missing price change for vacant units or aging since the calculation of implicit rent already adjusts for these considerations.

BLS makes 1-month and 6-month estimates of change in the owners' shelter cost for each market basket as follows: Let S_1 be the set of owner units with implicit rents in both time t and time t- 1 and S_6 the set of owner units with implicit rents in and t-6. Note that owner units may not be in S_1 or S6 if their sets Q_{j1} or Q_{j6} are empty. The 1-month and 6-month estimates of price change for owner units in each market basket are:

$$R_{t,t-1} = \frac{\sum_{j \varepsilon S_1} m_{jt} w_{j1}}{\sum_{j \varepsilon S_1} m_{j,t-1} w_{j1}} \qquad R_{i,t-6} = \frac{\sum_{i \varepsilon S_6} m_{jt} w_{j6}}{\sum_{i \varepsilon S_6} m_{j,t-6} w_{j6}}$$

[23]For additional information on quality adjustments in housing, see Henderson, Steve W. and Berenson, Stephen A. "Quality Adjustments for Structural Changes in CPI Housing Sample,'' *Monthly Labor Review,* November 1990, pp. 40-42.

[24] Substantial changes in the method of measuring price change of owner-occupied housing were introduced with the index for January 1983 (January 1985 for CPI-W). For information on the changes and the old method, see "changing the Homeownership Component of the Consumer Price Index to Rental Equivalence," CPI Detailed Report, January 1983.

where w_{j1} and w_{j6} are the owner units' inverse probability of selection adjusted for nonresponse.

As in the rent index, two preliminary estimates of the current month's price index for each market basket ($I_{t-1}R_{t,t-1}$ and $I_{t-6}R_{t,t-6}$) are averaged together to get the final estimate It:

$$It = A\ (It\text{-}1\ Rt,t\text{-}1) + (1\text{-}A)\ (It\text{-}6Rt,t\text{-}6)$$

Again, A = 0.65 is a value that reduces the mean squared error.

Estimation of seasonal effects

Seasonal adjustment. Seasonal adjustment removes the estimated effect of changes that normally occur at the same time every year (such as price movements resulting from changing climatic conditions, production cycles, model changeovers, holidays, sales, etc.). CPI series are selected for seasonal adjustment if they pass certain statistical criteria and if there is an economic rationale for the observed seasonality. Seasonal factors used in computing the seasonally adjusted indexes are derived using the ARIMA option of the X-11 variant of the Census Method II Seasonal Adjustment Program. In some cases, intervention analysis seasonal adjustment is carried out using X-12-ARIMA to derive more accurate seasonal factors. Consumer price indexes may be adjusted directly or aggregatively depending on the level of aggregation of the index, and the behavior of the component series. [25]

Intervention analysis seasonal adjustment. Some index series that show occasional erratic behavior known as a 'trend shift," which can cause problems in making an accurate seasonal adjustment. An index series whose underlying trend has experienced a sharp and permanent shift will generate distorted results when put through the X-11 ARIMA procedure. Trend shifts have been observed, for example, when gasoline prices have reacted to major changes instituted by the Organization of Petroleum Exporting Countries cartel—a recurring event which happens at infrequent and irregular intervals. Another kind of distorting change may occur when the seasonal pattern itself changes.

In order to compensate for those instances in which such distortions (called interventions) are both substantial and identifiable, regression techniques are used to model the distortiions and account for them as part of the seasonal adjustment process. Intervention analysis seasonal adjustment is performed using X-12-ARIMA seasonal adjustment software. X-12-ARIMA, developed by the Bureau of the Census, is an extension of the X-11 methodology which allows the use of regression-ARIMA models for more sophisticated time series analysis.[26] In recent years, BLS has used intervention analysis seasonal adjustment for various indexes—gasoline, fuel oil, new vehicle, women's apparel, and tobacco and smoking products.

Direct and aggregative adjustment. Each year BLS seasonally adjusts eligible lower level CPI index series independently with the X-11-ARIMA multiplicative model on to data for the latest 5 to 8 calendar years. This product's seasonal factors that will be used to generate seasonal adjusted indexes for the current year. BLS recalculates and publishes seasonally adjusted indexes for the previous 5 years.

Most higher level index series are adjusted by the indirect, or aggregative, method, which is more appropriate for broad categories whose component indexes show strongly different seasonal patterns. Under the aggregative method, direct adjustment is first applied to indexes at lower levels of detail, and thereafter the adjusted detail is aggregated up to yield the higher level seasonally adjusted indexes. If intervention analysis is indicated, it will be used in adjusting selected lower level indexes prior to aggregation. For those series which have not been selected for seasonal adjustment, the original, unadjusted data are used in the aggregation process.

Revision. The seasonal factors are updated annually. BLS recalculates and publishes seasonally adjusted indexes for the previous 5 years.

Average Prices

Average prices are estimated from CPI data for selected food items, gasoline, utility (piped) gas, electricity, and fuel oil in order to support the research and analytic needs of CPI data users. For each food item, the average price for aspecified unit of size (i.e., pound, gallon, etc.) are published monthly for the U.S. city average and for the four regions - Northeast, Midwest (formerly the North Central), South, and West. The regional definitions are those of the Bureau of the Census. Metric equivalent sizes are noted as well.

Average prices for gasoline, utility (piped) gas, electricity, and fuel oil are published monthly for the U.S. city average, the 4 regions, the 4 population size classes, 13 of 16 region/size-class cross classifications, 4 population size classes, and the 15 largest index areas. For utility (piped) gas, average prices per therm, per 40 therms, and per 100 therms are published. For electricity, average prices per kilo-

[25] See appendix A for an explanation of BLS seasonal adjustment methods.

[26] Buszuwski, J.A. and Scott, S. (1988), "On the use of Intervention Analysis in Seasonal Adjustment," *Proceedings of the Business and Economics Section,* American Statistical Association.

watt-hour (kwh) and per 500 kwh are published. For fuel oil and gasoline, the average price per gallon is published. Average prices for popular grades of gasoline are published.

Price quotes for 40 therms and 100 therms of utility gas and for 500 kwh of electricity are collected in sample outlets for use in the average price programs only. Since they are for prespecified consumption amounts, they are not used in the CPI. All other price quotes used for average price estimation are regular CPI data.

With the exception of the 40 therms, 100 therms, and 500 kwh price quotes, all prices are converted to a price per-normalized quantity. For example, prices for gallons, quarts, or pints of milk are converted to prices per ounce. All prices are then used to estimate a price for a defined fixed quantity. That is, a price per ounce of milk is estimated and multiplied by 64 to yield a price per half gallon, the published quantity.

The average price for collection period t is estimated as:

$$\overline{P}_t = \frac{\sum_i W_{it} P_{it} / P_{ia}}{\sum_i W_{it} / P_{ia}}$$

where W_{it} is the quote weight as defined in the estimation of price change modified to reflect the number of quotes usable for average price estimation for the ELI/PSU/replicate. (Imputed prices are used in estimating average prices.)

In the equation, W_{it} is an expenditure weight. Dividing the expenditure weight by the price, $P_{ia,}$ for a given quote yields an implicit estimate of quantity. Thus, the average price is, conceptually, a weighted average of prices where the weights are quantity amounts.

Part III. Precision of Estimates

An important advantage of probability sampling methods is that a measure of the sampling error of survey estimates can be computed directly from the sample data. The CPI sample design accommodates error estimation by making two or more selections (replicates) of items and outlets within an index area. Therefore, two or more samples of quotes in each self-representing PSU and one in each non-self-representing PSU are available. With this structure, which reflects all stages of the sample design, variance estimation techniques using replicated samples can be used.

Different methods of variance estimation were used 'for the CPI during the period 1978-86 than are used in the current CPI. The sampling of areas, outlets, and items for the CPI for both periods followed replicated sample designs. The 1982-84 CE Survey also employed a replicated sample design. However, the 1972-73 CE Survey, which provided the expenditure weights for the 1978-86 CPI, did not use a replicated sample design. Thus it was necessary to perform two additional steps in total variance estimation for 1978-86 to include the contribution to the variance from the 1972-73 CE that were not necessary in estimating total sampling variance for 1987 forward. Both methods are described below.

1978-86 CPI Variance Estimation

The method used to estimate CPI variances for geographic aggregates for 1978-86 depends upon the statistical independence of the estimated indexes for individual index areas. The independence assumption does not hold entirely for the 1978-86 time period because during that time period, the same item sample of ELI's was used in more than one index area within a region. However, since the selection of specific items to be priced involved considerable subsampling of the ELI's within outlets, the local index area indexes were regarded as statistically independent for variance estimation. For each index area, two independent estimates of the index were constructed using the replicates specified in the design. This required calculation of price relatives by replicate for each item stratum for each time period, application of the replicate relatives to the previous-period replicate index for each item stratum, and aggregation across item strata to produce replicate indexes for item aggregates in each index area.

Squared differences of these indexes (properly scaled) provided preliminary estimates of the variance of the area index. These variance estimates were conditional on the values of the December 1977 expenditures which were updated from estimates from the 1972-73 CE Survey. Unconditional estimates of index variances, which include a component attributable to variation in expenditure weight estimates, were then computed using the conditional index variances and December 1977 expenditure weight variances which were estimated separately. Variance estimates were computed for all items and for major groups and selected item subcategories.

As noted earlier, in CPI estimation price relatives are computed for each item stratum for each index area of the CPI. Variance estimation required that price relatives also be computed for each item stratum for each index area and replicate. The methodology for computing the price relatives was the same for the full index area as for the replicates. All replicate computations were for the CPI-U population with sales and excise taxes included.

For commodities and services, each replicate sample was selected independently within each index area. In the shelter survey, for each self-representing PSU, each rental unit was assigned to one of two replicates. For non-self-representing index areas, the index PSU determined the replicate for a given rental unit.

For relative computation for rent and owners' equivalent rent, preliminary replicate cost weights for each index area replicate were constructed to provide a basis for weighting the 1-month and 6-month relatives together. For replicate j, item stratum i (rent or owners' equivalent rent), in index area m, let:

$r_j(i,m,t,t-1)$	denote the corresponding 1-month relative between months t and t-1,
$CW_j(i,m,t)$	denote the cost weight at month t, and
$r_j(i,m,t,t-6)$	denote the corresponding 6-month relative between months t and t-6.

Then $CW_j(i,m,t)$ is computed by:

$$CW_j(i,m,t) = 0.65\, CW_j(i,m,t-1)\, r_j(i,m,t,t-1) + 0.35\, CW_j(i,m,t-6)\, r_j(i,m,t,t-6)$$

The final shelter relative, $R_j(i,m,t,t-1)$, is computed by:

$$R_j(i,m,t,t-1) = CW_j(i\ m\ t) / CW_j(i,m,t-1)$$

The CPI for any item stratum in any index area is computed by a chaining process in which the estimate of expenditure for the previous month for the item stratum is multiplied by its 1-month price relative to provide an estimate of the current month's expenditure for the stratum. Item stra-

tum expenditure values, called cost weights and denoted by CW(i,m,t) are then aggregated and compared to the total expenditure in the reference period denoted by 0. Thus, for a single item stratum index the sequence of computations would be:

$$CW(i,m,t) = CW(i,m,t-1) \cdot R(i,m,t,t-1)$$

and

$$IX(i,m,t) = 100 \cdot [CW(i,m,t) / CW(i,m,0)].$$

Estimating Variances of the Index and Price Change

Estimates of the variance of the index, conditional on the December 1977 expenditure weights, for all items or a subset of items at the national, regional, or area level, were calculated using a random group estimation method. As noted above, the sample for the CPI in any index area is partitioned into two or more disjoint replicate panels, also termed random groups. In self-representing index areas comprising only one PSU, these replicate panels are disjoint subsets of the sample for the PSU. The sample for most self-representing index areas consists of two replicates. In non-self-representing index areas, each replicate consists of the sample for one or more of the sample PSUs in the index area. The number of replicates for non-self-representing index areas in this time period ranged from two to four.

Beginning in January 1978, price relatives for each item stratum were computed separately for the full sample and for each replicate in every CPI index area to produce separate cost weight series for all item strata for each replicate as well as the full sample for each area.

Cost weights for higher level item aggregates were constructed for each index area at full sample and replicate levels using the index aggregation methodology described earlier. Item stratum and item aggregate index variance estimates were then computed for each index area using the following formulas:

$$Var_c[IX(I,m,t)] = \frac{\sum_{j=1}^{r}\left[CW_j(I,m,t) - CW_f(I,m,t)\right]^2}{r(r-1)\,CW(I,m,0)^2}$$

where I denotes the item or item aggregate, r denotes the total number of replicates in the index area, and $CW_f(I,m,t)$ and $CW_j(I,m,t)$ denote the cost weights for the full sample and replicate j, respectively, in the index area m.

Conditional mixed index covariances between item strata and I and periods t and t-k were estimated analogously:

$$Cov_c[IX(i,m,t), IX(1,m,t-k)] = \frac{\sum_{j=1}^{r}\left[CW_j(i,m,t) - CW_f(i,m,t)\right]\left[CW_j(1,m,t-k) - CW_f(1,m,t-k)\right]}{r(r-1)CW(i,m,0)\left[CW(1,m,0)\right]}$$

The December 1977 cost weights for the full sample and all replicates for any series in any index area were equal in value. Variances computed using these cost weight series alone are termed conditional because they do not reflect the variance of the index or price change due to sampling variation of the December 1977 expenditure estimates. That is, they are conditional on the values of December 1977 expenditure estimates. Also for 1978-86, base period expenditure estimates were not separately estimated for full and replicate samples, and thus were not distinguished.

In addition, for the database from which these variances were estimated, cost weights for the Homeowners Equivalent Rent (REQ) item stratum were computed for every month in the 9-year period, even though the REQ stratum was not officially incorporated in the index before January 1983. These REQ cost weights were used in computation of the index and price change and their variances for All Items and Housing for each month. Thus the estimates of the index and price change for All Items, Housing, and Shelter used in variance estimation for 1978-82 do not correspond exactly with their published values.

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 1972-73. Expenditure data for REQ were obtained from the 1974 CE surveys. Each consumer unit was assigned a full sample weight and a set of 36 replicate weights. For each replicate, region, index area, and stratum the estimation of final mean expenditures was accomplished using three steps:

(1) preliminary estimation of expenditures,
(2) estimation of mean expenditures using a composite estimation procedure which combined index area and regional expenditure estimates,
(3) estimation of final mean expenditures using a raking process on the mean expenditures determined in (2).

For each item-area, the final raked mean expenditures were combined from the Diary and Quarterly Surveys. Final raked mean expenditures were then adjusted to reflect the new item strata structure which was revised in 1977. This was done by applying ratio adjustment factors to obtain mean expenditure estimates for newly introduced item strata. Inflation factors were then applied to estimate expenditures for December 1977, denoted 7712 below.

Cost weight variances for 7712 were estimated for each index area-stratum combination by:

$$Var[CW(i,m,7712)] = \left[\sum_{j=1}^{36}\left(RME_j(i) - RME_f(i)\right)^2\right] \Big/ 36$$

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 an estimate of final raked full sample mean ex-

penditure. Similarly, between item strata covariances were estimated by:

$$\mathrm{Cov}\left[CW(i,m,7712),\ CW(1,m,7712)\right] = \left\{\sum_{j=1}^{36}\left[RME_{\ j}(i,m,7712) - RME_{\ f}(i,m,7712)\right]\cdot \left[RME_{\ j}(1,m,7712) - RME_{\ f}(1,m,7712)\right]\right\} \Big/ 36$$

Cost weight variance estimates were then combined with conditional variances to produce unconditional estimates of index and price change variances. The unconditional variance of the index for time t was estimated by:

$$\mathrm{Var}\left[IX(I,m,t)\right] = \left[100\ /\ CW(I,m,0)\right]^2 \cdot \left\{\mathrm{Var}\left[CW(I,m,t)\right] + \left[CW(I,m,t)/CW(I,m,7712)\right]^2 \cdot \mathrm{Var}\left[CW(I,m,7712)\right] - 2\cdot\left[\left(CW\langle I,m,t\rangle\right)/CW(I,m,7712)\right] \cdot \mathrm{Cov}\left[CW(I,m,t),\ CW(I,m,7712)\right]\right\}$$

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

$$PC(I,m,t,t-k) = 100\cdot\left[\left(CW\langle I,m,t\rangle / CW\langle I,m,t-k\rangle\right) - 1\right],$$

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

$$\mathrm{Var}\left[PC(I,m,t,t-k)\right] = \left[100\ /\ CW(I,m,t-k)\right]^2 \cdot \left\{\mathrm{Var}\left[CW(I,m,t)\right] + \left[CW(I,m,t)/CW(I,m,t-k)\right]^2 \cdot \mathrm{Var}\left[CW(I,m,t-k)\right] - 2\cdot\left[CW(I,m,t)/CW(I,m,t-k)\right]\mathrm{Cov}\left[CW(I,m,t),CW(I,m,t-k)\right]\right\}$$

Here Var[CW(I,m,7712)j 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 item or item aggregate I, index area m, and times t and t-k, estimated via the equality:

$$\mathrm{Var}\left[CW(I,m,t)\right] = \mathrm{Var}\left\{E\left[CW(I,m,t)\middle|CW(I,m,7712)\right]\right\} + E\left\{\mathrm{Var}\left[CW(I,m,t)\middle|CW(I,m,7712)\right]\right\}$$

which gives:

$$\mathrm{Var}\left[CW(I,m,t)\right] = \left[CW(I,m,o)/100\right]^2 \cdot \mathrm{Var}_{\,c}\left[IX(I,m,t)\right] + \sum_{i\varepsilon I}\sum_{j\varepsilon I}\left\{\left[IX(i,m,7712)\cdot IX(j,m,7712)\right]^{-1} \cdot \left[IX(i,m,t)\right]\cdot\left[IX(j,m,t)\right] - \mathrm{Cov}_{\,c}\left[IX(i,m,t),IX(j,m,t)\right] \cdot \mathrm{Cov}\left[CW(i,m,7712),CW(j,m,7712)\right]\right\}$$

and

$$\mathrm{Cov}\left[CW(I,m,t),CW(I,m,t-k)\right] = \left[CW(I,m,0)/100\right]^2 \cdot \mathrm{Cov}_{\,c}\left[IX(I,m,t),IX(I,m,t-k)\right] + \sum_{i\varepsilon I}\sum_{j\varepsilon I}\left\{\left[IX(i,m,7712)\cdot IX(j,m,7712)\right]^{-1} \cdot IX(i,m,t)\cdot IX(j,m,t-k) - \mathrm{Cov}_{\,c}\left[IX(i,m,t),IX(j,m,t-k)\right] \cdot \mathrm{Cov}\left[CW(i,m,7712),CW(j,m,7712)\right]\right\},$$

Index and price change total variance estimates for higher level geographic aggregates were computed, assuming independence of indexes between areas, by summing cost weight variances and covariances over index areas within the geographic aggregate M:

$$\mathrm{Var}\left[IX(I,M,t)\right] = \left[100\ /\ CW(I,M,0)\right]^2 \cdot \left\{\mathrm{Var}\left[CW(I,M,t)\right] + \left[CW(I,M,t)/CW(I,M,7712)\right]^2 \cdot \mathrm{Var}\left[CW(I,M,7712)\right] - 2\cdot\left[CW(I,M,t)/CW(I,M,7712)\right]\cdot \mathrm{Cov}\left[CW(I,M,t),CW(I,M,7712)\right]\right\},$$

and

$$\mathrm{Var}\left[PC(I,M,t,t-k)\right] = \left[100\ /\ CW(I,M,t-k)\right]^2 \left\{\mathrm{Var}\left[CW(I,M,t)\right] + \left[CW(I,M,t)/CW(I,M,t-k)\right]^2 \cdot \mathrm{Var}\left[CW(I,M,t-k)\right] - 2\cdot\left[CW(I,M,t)/CW(I,M,t-k)\right]\cdot \mathrm{Cov}\left[CW(I,M,t),CW(I,M,t-k)\right]\right\}$$

where:

$$\mathrm{Var}\left[CW(I,M,t)\right] = \sum_{m\,\varepsilon M}\mathrm{Var}\left[CW(I,m,t)\right] \quad \text{and}$$

$$\mathrm{Cov}\left[CW(I,M,t),CW(I,M,t-k)\right] = \sum_{m\,\varepsilon M}\mathrm{Cov}\left[CW(I,m,t),CW(I,m,t-k)\right].$$

Estimating Variances of the Index and of Price Change, Starting in 1987

There are two important differences between the variance estimates for the CPI series starting with the 1987 revision and those computed for 1978-86. The first difference between the estimates derives from the fact that the December 1986 revision expenditure estimates based on the 1982-84 Consumer Expenditure Survey were independently estimated for each replicate. Thus the variances computed for the 1987 revision index series directly incorporate the contribution to sampling variance attributable to the estimation of expenditure weights from the 1982-84 CE Survey, and so are unconditional estimators of the variance of the index or price change.

The second difference between the estimates is that the 1987 revised CPI variances incorporate between-index-area covariances in estimates for higher level geographic area aggregates such as regions, city-size classes, and All Cities.

As with the 1978-86 index series, BLS will be estimating variances for the index series for the CPI-U population with sales and excise taxes included. Expenditure weights used in revising item-area weights for the index for this period were derived from the 1982-84 Consumer Expenditure Survey.

As before, the variance estimators given here depend on the aggregation structure of the index which supports the construction of indexes for higher levels of aggregation such as item groups, regions, and All Cities from those for basic item strata and index areas. The estimators also depend on the replicate structure of the index sample. Use of replicates

provides a means of measuring the overall variation of the index from those computed over subsets of the sample. The full sample for each index area comprises two or more replicate panels, half of which were designated ''odd'' and the other half ''even.''

Each index area is in one of four Census regions. Each region can further be divided into two major areas, one composed of the self-representing (A) index areas and one composed of the non-self-representing (non-A) index areas. Hence, there are eight major areas in the Nation-For each area aggregate larger than one index area, estimates of between-index-area covariances for each pair of different index areas in the same major area in the same area aggregate were included in index and price change computation.

Variance estimates for the index

To estimate the unconditional variance of an index, consider the 2n x 1 vector $\mathbf{CW}_f(IMA,t,t')$ of full sample cost weights for an item or item aggregate I, whose elements are the cost weights for each of n index areas in major area MA in months t and t′:

$$\mathbf{CW}_f(I,MA,t,t') = [CW_f(I,m_1,t),...,CW_f(I,m_n,t),CW_f(I,m_1,t'),...,CW_f(I,mn)]^T$$

Similarly denote $\mathbf{CW}_1(I,MA,t,t')$ and $\mathbf{CW}_2(I,MA,t,t')$ to be the corresponding vectors of average replicate cost weights for the item and major area, with 1 denoting the average of r/2 odd replicates and 2 denoting the average of r/2 even replicates:

$$[\mathbf{CW}_1(I,MA,t,t') = 2/r \sum_{j\ odd} CWj(I,MA,t,t')$$

$$(\mathbf{CW}_2(I,MA,t,t') = 2/r) \sum_{j\ even} CWj(I,MA,t,t')$$

Let $^{A}_{M,MA}$ be the 2 x 2n area aggregation matrix for any area aggregate M with component index areas belonging to major area MA where:

$\mathbf{A}_{M,MA}(1,j)$ = 1 if index area jeM, 0 otherwise; j=1,...,n
$\mathbf{A}_{M,MA}(1,j)$ = 0; j=n+1,...,2n
$\mathbf{A}_{M,MA}(2,j)$ = 0; j=1,...,n, and
$^{A}_{M,MA}(2,j)$ = 1 if index area jeM, 0 otherwise; j=n+1,...,2n.

The covariance matrix of cost weights for any area aggregate M within a major area MA, $\mathbf{W}(I,M,MA,t,t')$, is estimated by:

$$(\mathbf{W}(I,M,MA,t,t') = \mathbf{A}_{M,MA}\ 1/2[\sum_{j=1}^{2} \mathbf{DCW}j(I,MA,t,t')\,\mathbf{DCW}j(I,MA,t,t')T]\ ^{\mathbf{A}}M,MAT$$

where $\mathbf{DCW}j(I,MA,t,t')$ is the difference vector:

$$\mathbf{DCW}_j(I,MA,t,t') = [\mathbf{CW_j}(I,MA,t,t') - \mathbf{CW}_f(I,MA,t,t')],\ j=1,\ 2.$$

Under the assumption that the cost weights for index areas are independent between major areas, the cost weight covariance matrix $\mathbf{W}(I,M,t,t')$ for any area aggregate M comprising index areas in more than one major area, such as All Cities, regions, and size classes, is computed by summing $\mathbf{W}(I,M,MA,t,t')$ over all major areas:

$$\mathbf{W}(I,M,t,t') = \sum_{MA=1}^{8} [\mathbf{W}(I,M,MA,t,t')]$$

Given the cost weight covariance matrix **W**, the variance of the index or price change can be estimated by a first order Taylor series approximation of the ratio of two cost weights at times t and t′ The exact expansions are given below.

For the variance of the index at time t, t′ is December 1986, denoted 8612. Let $\mathbf{L}(I,M,t,8612)$ represent 1 x 2 the linear transformation vector:

$$\mathbf{L}(I,M,t,8612)=IX(I,M,8612)\bullet\{1/CW(I,M,8612),\ -CW(I,M,t)/[CW(I,M,8612)]^2\}.$$

Then the variance of IX(i,m,t) is estimated by:

$$Var[IX(I,M,t)]=\mathbf{L}(I,M,t,8612)\mathbf{W}(I,M,t,8612)\,\mathbf{L}(I,M,t,8612)^T$$

Variance estimates for price change

An estimate of k-month price change, PC(I,M,t,t-k), from month t-k to month t for item aggregate I and area aggregate M is computed by:

$$PC(I,M,t,t\text{-}k) = 100 \bullet \{[1CW(I,M,t)/CW(I,M,t\text{-}k)]\text{-}1)\}.$$

Thus, price change is also a simple function of the ratio of cost weights for two time periods. Its variance can be estimated by:

$$Var[PC(I,M,t,t\text{-}k)] = \mathbf{L}(I,M,t,t\text{-}k)\ \mathbf{W}(I,M,t,t\text{-}k)\ \mathbf{L}(I,M,t,t\text{-}k)^T$$

where $\mathbf{W}(I,M,t,t\text{-}k)$ is as defined above and the linear transformation vector $\mathbf{L}(I,M,t,t\text{-}k)$ is given by:

$$\mathbf{L}(I,M,t,t\text{-}k) = 100\bullet\{1/CW(I,M,t\text{-}k),\ -\ CW(I,M,t)/[CW(I,M,t\text{-}k)]^2\}.$$

Nonsampling Error

CPI estimates are subject to nonsampling error as well as sampling error. Surveys involve many operations that must

be performed in order to produce the final results. All of these are potential sources of nonsampling error. The errors arise from the survey process regardless of whether the data are collected from the entire universe or from a sample of the population. The most general categories of nonsampling error are coverage error, nonresponse error, response error, processing error, and estimation error.

Coverage error in an estimate results from the omission of part of the target population (undercoverage) or the inclusion of units from outside of the target population (overcoverage). Coverage errors result from the omission of cities, households, outlets, and items that are part of the target populations from the relevant sampling frames or from the double counting or inclusion of them in the frames when they should not be. A potential source of coverage error is the time lag between the Point-of-Purchase Survey and the initiation of price collection for commodities and services at resampled outlets. Because of the time lag, the products offered by the outlet at the time pricing is initiated may not coincide with the set from which the POPS respondents were purchasing.

Nonresponse error results when data are not collected for some sampled units because of the failure to interview households or outlets. This can occur when selected households and outlets cannot be contacted or refuse to participate in the survey. Nonresponse rates at initiation for the GPI commodities and services and housing surveys are shown in table 3. This nonresponse could bias the CPI if the rate of price change at the nonresponding survey units differed from the rate of price change at the survey units successfully initiated. Nonresponse rates during monthly pricing for the CPI commodities and services and housing surveys are shown in tables 1 and 2.

Response error results from the collection and use in estimation of incorrect, inconsistent, or incomplete data. Response error may arise because of the collection of data from inappropriate respondents, respondent memory or recall errors, deliberate distortion of responses, interviewer effects, misrecording of responses, pricing of wrong items, misunderstanding or misapplication of data collection procedures, or misunderstanding of the survey needs and/or lack of cooperation from respondents. The pricing methodology in the commodities and services component of the CPI allows the previous period's price to be available at the time of collection. This dependent pricing methodology is believed to reduce response variance for measuring change, but may cause response bias and lag. The housing component of the CPI, however, employs an independent pricing methodology specifically to avoid potential response bias.

In 1986 BLS established a program called Process Audit to investigate various kinds of response variance in the CPI. Process auditors independently recollect pricing information for a subsample of outlets and ELI's drawn from the regular commodities and services sample. The primary objective has been to determine where and to what extent discrepancies occur between the Process Audit collections and the routine submissions from the field. Discrepancies are not just errors but also contain actual price changes that occur in the brief time period between field and process audit collections and cannot be easily identified. Studies have revealed that these pricing discrepancies tend to be related to certain pricing characteristics of the item, such as the amount of interaction required between the data collector and the respondent in order to collect the data, those items that have multiple components to be priced, the length of the checklist that is used to fully describe the item, and the degree of price volatility of the item. A recent redesign of the Process Audit sample will allow for a tracking of these discrepancy rates from year to year making it possible to measure the impact of process improvements and to identify new or existing problems in the pricing process.

Table 1. Response rates for commodities and services for the CPI-U, U.S. City Average, by major group, 1990

Commodities and services	Eligible	Collected	Percent collected	Used in estimation	Percent in estimation
Outlets	240,212	231 613	96.4	222,264	92.5
Total quotes	962,322	845,657	87.9	826,966	86.1
Food	526,600	484,871	92.1	477,945	90.8
Housing (excluding shelter)	119,725	111,799	53.1	109,212	91.2
Apparel and upkeep	131,547	79,465	60.4	75,181	57.2
Transportation	73,062	67,805	92.8	66,799	91.4
Medical care	44,623	41,118	92.8	40,833	91.5
Entertainment	37,363	31,644	84.7	30,630	82.0
Other goods and services	30,507	28,685	91.0	28,368	93.1

Table 2. Response rates for housing for the CPI-U, U.S. City Average, 1990

Housing (Shelter)	Eligible	Collected, data reported	Collected, found vacant	No Data at collection or other	Used in estimation
Number of total units	107,059	82,643	6,699	17,717	82,465
Percent of eligible units	100.0	77.2	6.3	16.5	77.0
Number of owners	31,664	26,945	838	3,881	27,069
Percent of eligible owners	100.0	85.1	2.6	12.3	85.5
Number of renters	69,900	55,569	5,636	8,695	55,396
Percent of eligible renters	100.0	79.5	8.1	12.4	79.3
Number of unknown	5,495	129	225	5,'41	0
Percent of eligible unknown	100.0	2.3	4.1	93.6	0.0

Table 3. Response rates at initiation of commodities and services for the CPI-U, U.S. City Average, 1990

Commodities and services	Eligible	Collected	Percent collected
Outlets	3,620	3,403	89.1
Quotes	16,360	12,080	73.8

Processing error arises from incorrect editing, coding, and data transfer. Survey data are converted into machine-readable form by two independent key entry operators, and discrepancies are resolved by a third person. Processing errors can be introduced by an incorrect resolution or by an identical miskeying of an element by two operators. Errors can also result from software problems in the computer processing which cause correctly keyed data to be lost. Computer screening and professional review of the data provide checks on processing accuracy. Occasional studies of these processing errors in the CPI have shown them to be extremely small.

Estimation error results when the survey process does not accurately measure what is intended. Such errors may be conceptual or procedural in nature, arising from a misunderstanding of the underlying survey measurement concepts or a misapplication of rules and procedures. A source of estimation error due to conceptual problems was the treatment of housing before 1983, which failed to distinguish between the consumption and investment aspects of homeownership. Prior to implementation of the change to the owners' equivalent rent, an experimental measure using rental equivalence diverged considerably from the official CPI.[27]

Substitutions and adjustments for quality change in the items priced for the CPI are possible sources of estimation error due to procedural difficulties. Ideally, CPI data collection forms and procedures would yield all information necessary to determine or explain price and quality differences for all items defined within an ELI. Since such perfect information is not available, BLS economists supplement directly collected data with secondary data. Estimation error will result if the BLS adjustment process, which may require significant judgment or lack key data, is misapplied, or if it consistently overestimates or under-estimates quality change for particular kinds of items. While individual problems arising from estimating quality change have been identified, the evidence to date is that on average there is no systematic bias from this process. Cases where price change is overestimated are about as frequent as those where it is underestimated.

The effect of the aging of housing units is an example of potential estimation error, which is similar to the issue of quality change in commodities and services. Until 1988, BLS did not adjust for the slow depreciation of houses and apartments over time. Current BLS research indicates that annual changes for the residential rent and owners' equivalent rent indexes would have been 0.1 to 0.2 percent larger if some type of aging adjustment had been included.

The total nonsampling error of the CPI results from errors in the type of data collected, the methods of collection, the data processing routines, and the estimation processes. The cumulative nonsampling error can be much greater than the sampling error.

Response rates

Response rates are calculated for the CPI-U at the data collection phase and at the index estimation phase for ongoing pricing. The response rate at the data collection phase is the number of responding sample units divided by the sum of (1) the number of eligible sample units and, (2) the number of sample units with eligibility not determined. A sample unit is eligible if it belongs to the defined target population and responses should be collected from the unit for one or more items. The response rate at estimation is defined as the number of sample units used in estimation divided by the sum of (1) the number of eligible sample units and, (2) the number of sample units with eligibility not determined.

Commodities and services items (any except rent and owner's equivalent rent) are further broken down into outlets and quotes. An "outlet" is a generic term used by the CPI to describe places where prices are collected. A "quote" is a specific item to be priced in a specific outlet. There may be from 1 to more than 50 quotes priced in an outlet. In table 1, it is important to note the relatively low percentages of quotes reported collected and used in estimation for apparel and upkeep. Low rates for these items can be largely attributed to the design of the apparel sample. Because apparel items are commonly in stores only during certain times of the year, most of the apparel sample is doubled, with each half of the sample designated for pricing during part of the year. Thus, at any particular time of the year a large number of apparel quotes, although eligible, are designated out of season and prices are not collected. For further information see the section above on seasonal items.

The response rates for housing (shelter) shown in table 2 include categories for owners, renters, and those for whom the tenure status is unknown. A unit qualifies as an owner or renter if its tenure status is known either by previous knowledge or is collected in the current interview period. The number of responses is higher than the total sample size because some of the sample is priced twice a year. The response rates at the data collection phase for housing (shelter) are separated into three categories. If usable information is obtained, the unit is designated collected, data reported. If the assigned unit is located but is unoccupied, the unit is designated collected, found vacant. In instances where the unit is eligible but no data are available (e.g. refusals), the unit is designated no data at collection or other.

[27] See Gillingham, Robert and Lane, Walter. "Changing the Treatment of Shelter Costs for Homeowners in the Consumer Price Index," *Statistical Reporter*, December 1981.

Data for initiation of commodities and services (table 3) are for those areas which underwent sample rotation in 1990. Approximately one-fifth of the areas (primary sample units) are rotated each year. The response rate data for initiation reflect the rate of success in attempting to price the new designated sample for the first time. These data are unique in that the outlets and quotes are only counted once, as opposed to the repeated (monthly or bimonthly) pricing that is reflected in table1.

Technical References

Allen, R.G.D. *Index Numbers in Theory and Practice.* pp. 2-9.

Early, John F. and Galvin, John M. "Comprehensive Quality Measurement in the Consumer Price Index," *Proceedings of the 1987 Juran Institute Inc*., Annual Quality Management Conference.

Early, John F. and Sinclair, James H. "Quality Adjustment in the Producer Price Indexes," *The U.S. National Income and Product Accounts Selected Topics*, Murray F. Foss, ed . University of Chicago Press, 1983.

Effort, S. *The Price Index.* Cambridge University Press, 1977.

Fay, Robert E., *VPLX: Variance Estimates for Complex Samples*, Washington, DC: U.S. Bureau of the Census, 1993.

Ford, Ina K., and Sturm, Phillip, "CPI Revision Provides more Accuracy in the Medical Care Services Component," *Monthly Labor Review,* April 1988, pp. 17-26.

Frisch, R. "Annual Survey of Economic Theory: The Problem of Index Numbers," *Econometrica*, Vol. 4, January 1936.

Ghaemi A. and McCormick, G. P. ,"Factorable SUMT: What Is It? How is It Used?" Technical Paper Serial T-402, Washington, D.C., The George Washington University, Institute for Management Science and Engineering, 1979.

Gillingham, Robert F. "Estimating the User Cost of Owner-Occupied Housing," *Monthly Labor Review,* February 1980, pp. 31-35.

Gillingham, Robert F. "Measuring the Cost of Shelter for Homeowners: Theoretical and Empirical Considerations," *The Review of Economics and Statistics, 1983*, pp. 254-65.

Gillingham, Robert F. and Greenlees, John S. "The Impact of Direct Taxes on the Cost of Living," *Journal of Political Economy*, 95(4), August 1987.

Gillingham, Robert F. and Lane, Walter. "Changing the Treatment of Shelter Costs for Homeowners in the CPI," *Monthly Labor Review,* June 1982, pp. 9-14.

Grandits, Steven. "1998 CPI Revision:Publication Strategy," *Monthly Labor Review,* December 1996.

Greenlees, John S. "1998 CPI Revision: Revison Overview," *Monthly Labor Review,* December 1996.

Hageman, Robert F. "The Variability of Inflation Rates Across Household Types," *Journal Of Money, Credit and Banking*, Vol. 14:4, 1982, pp. 494-510.

Kokoski, Mary F. "Consumer Price Indexes by Demographic Group," BLS Working Paper 167, April 1987.

Kokoski, Mary F. *Experimental Cost-of-Living Index Measures,* BLS Report 751, June 1988

Kosary, Carol L. and Sommers, John F. "The Effect of Computer Assisted Telephone Interviewing on Response Inconsistency," *Proceedings of the Survey Research Methods Section*, American Statistical Association, 1987, pp. 627-632.

Kosary, Carol L.; Sommers, John F.; and Branscome, James M. "Evaluating Alternatives to the Rent Estimator," *Proceedings Of the Business and Economic Statistics Section*, American Statistical Association, 1984, pp. 410-12.

Lane, Walter F. "1998 CPI Revision: Item Structure," *Monthly Labor Review,* December 1996.

Lane, Walter F. and Sommers, John F. "Improved Measures of Shelter Costs," *Proceedings Of the Business and Economic Statistics Section*, American Statistical Association, 1984, pp. 49-56.

Leaver, Sylvia G. "Estimating Conditional Variances of the U.S. Consumer Price Index for 1978-86," *Proceedings of the Survey Research Methods Section*, American Statistical Association, 1990, pp. 290-295.

Leaver, Sylvia G., Johnstone, James E., and Archer, Ken neth P., "Estimating Unconditional Variances of the U.S. Consumer Price Index for 1978-86," *Proceedings of the Survey Research Methods Section*, American Statistical Association, 1991.

Leaver, Sylvia G., and Swanson, David C., "Estimating Variances for the U.S. Consumer Price Index for 1987-1991," *Proceedings of the Survey Research Methods Section*, American Statistical Association, 1992, pp. 740-745.

Leaver, Sylvia and Richard Valliant, "Statistical Problems in Estimating the U.S. Consumer Price Index," chapter 28, *Business Survey Methods.* Wiley & Sons, Inc., 1995.

Leaver, Sylvia G.; Weber, William L.; Cohen, Michael P.; and Archer, Kenneth P. "Item-Out-let Sample Redesign for the 1987 U.S. Consumer Price Index Revision," *Proceedings of the 46th Session, International Statistical Institute.* Tokyo, 1987, Vol. 3, pp. 173-185.

Liegey Jr., Paul R. "Adjusting Apparel Indexes in the CPI for Quality Differences," *Price Measurements and their Uses*, Murray F. Foss, Marilyn Manser, and Allan Young, eds. NBER-University of Chicago Press 1992.

Mason, Charles and Cage, Robert. "1998 CPI Revision: Item Rotation," *Monthly Labor Review,* December 1996.

Pollack, Robert A. *The Theory of the Cost-of-Living Index.* Oxford University Press, 1989.

Primont, Diane F., and Kokoski, Mary F., "*Comparing Prices Across Cities: A Hedonic Approach,"* BLS Working Paper 204, May, 1990.

Ptacek, Frank and Baskin, Robert. "1998 CPI Revision: Housing," *Monthly Labor Review*, December 1996.

Randolph, William C. "Estimation of Housing Depreciation: Long-term Vintage Effects," *Journal of Urban Economics*, Vol. 23, March 1988, pp. 162-178.

Randolph, William C. "Housing Depreciation and Aging Bias in the Consumer Price Index," BLS Working Paper 166, April 1987.

Randolph, William C. and Zieschang, Kimberly D. "Aggregation Consistent Restriction-Based Improvement of Local Area Estimation," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 1987.

Sommers, John P. and Rivers, Joseph D. "Vacancy Imputation Methodology for Rent, in the CPI," *Proceeding Of the Business, and Economic Statistics Section*, American Statistical Association, 1983, pp. 201-05.

Triplet, Jack E. "Reconciling the CPI and PCE Deflator," *Monthly Labor Review*, September 1981, pp. 3-15.

Williams, Janet L. "1998 CPI Revision: Area Design," *Monthly Labor Review*, December 1996.

Williams, Janet L., Brown, Eugene F., Zion, Gary R. "The Challenge of Redesigning the Consumer Price Index Area Sample," *Proceedings of the Survey Research Methods Section , American Statistical Association* (Vol. 1), 1993, pp. 200-205.