Consumer Price Index

The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by urban consumers for a representative basket of consumer goods and services. The CPI measures inflation as experienced by consumers in their day-to-day living expenses. Indexes are available for the United States and various geographic areas. Average price data for select utility, automotive fuel, and food items are also available. CPI indexes are used to adjust income eligibility levels for government assistance, federal tax brackets, federally mandated cost-of-living increases, private sector wage and salary increases, poverty measures, and consumer and commercial rent escalations. Consequently, the CPI directly affects hundreds of millions of Americans.

.

Quick Facts: Consumer Price Index	
Subject areas	Prices
Key measures	Average prices Consumer price indexes Consumer prices
How the data are obtained	Survey of businesses, Survey of households
Classification system	Commodity
Periodicity of data availability	Bimonthly, Monthly
Geographic detail	Census region, Metro area, U.S. City average
Scope	Urban consumers
Key products	• Consumer Price Index news release • Consumer Price Index factsheets • Databases • Tables • Interactive charts
Program webpage	www.bls.gov/cpi

Table of Contents

CONCEPTS 2
DATA SOURCES.....6
DESIGN.....10
CALCULATION.....18
PRESENTATION 54
HISTORY.....58
MORE INFORMATION....63

Concepts

The Consumer Price Index (CPI) consists of a family of indexes that measure price change experienced by urban consumers. Specifically, the CPI measures the average change in price over time of a market basket of consumer goods and services. The market basket includes everything from food items to automobiles to rent. The CPI is perhaps the most noted measure of consumer inflation in the United States, and it is used by policy makers to understand and analyze the economy. It is used in many official contexts, for example to escalate Social Security and other federal payments, to adjust tax brackets, to deflate other time series data, and to convert nominal dollars to real dollars. It is also widely used by businesses and private citizens to adjust wages and to escalate rents and other payments.

The CPI as a cost-of-living framework

The CPI is widely used as a cost-of-living index, which answers the hypothetical question concerning what expenditure level is needed to achieve a standard of living attained in a base period at current market prices. The ratio of this hypothetical cost to the actual cost of the base-period consumption basket in the base period is the cost-of-living index. The cost of living is affected by many things not captured in market transactions, and the cost of achieving a living standard cannot be observed directly, so the CPI only approximates a cost-of-living index. The CPI is sometimes called a conditional cost-of-living index, since the factors that affect the cost of living that aren't in scope are implicitly held constant. The concept of the cost-of-living index guides the CPI measurement objective and is the standard by which any bias in the CPI is defined.

The CPI is constructed using a set of surveys, and it is fundamentally a measure of price change. The CPI follows the prices of a sample of items in various categories of consumer spending, encompassing a majority of all goods and services purchased by urban consumers for consumption.

The CPI focuses on the consumer experience of inflation, therefore the price sought is typically the consumer's out-of-pocket price, including sales and excise taxes. This contrasts with the Producer Price Index, which focuses on what is received by the producer. Prices to be used in the estimation of the CPI are collected during the course of the entire month, which is subdivided into three pricing periods, with a portion of the sample assigned to each. CPI data correspond to a month, not a specific date.

The monthly movement in the CPI comes from weighted averages of the price changes of the items in the CPI sample. A sample item’s price change is the ratio of its price at the current time to its price in a previous time.

CPI index values

The CPI computes and publishes index values, which are normalized to equal 100 in a chosen base period (1982–84 for most indexes). Index values can be interpreted as representing an estimate of the price level relative to the base period. Percent change in the index is an estimate of the percent change in the price level over the period in question. The CPI publishes index values, along with 1-month and 12-month changes; the 12-month change is the most frequently referenced change. Note that while index values serve as a proxy for the price level, index levels

of different series cannot meaningfully be compared to each other. That is to say, an index level of 200 for one region or metropolitan area does not represent a higher price level than a value of 150 for another area.

While the CPI program primarily publishes indexes that measure the average change in price over time, it also publishes average price data for a limited set of food and energy items for the United States, Census Regions, Census Divisions, and published metropolitan areas.

Structure and classification

The CPI is calculated in a two-stage process. First, basic indexes are calculated; these are indexes for specific item-area combinations. Ice cream and related products in the Chicago-Naperville-Elgin metro area are an example. These are structured by item category and geographic location. In the second stage, the basic indexes are aggregated into broader indexes, all the way up to the all items U.S. city average index. Thus, the CPI has both a geographic structure and an item structure.

Items

Expenditure items are classified in the CPI into more than 200 categories, arranged into 8 major groups. This item structure is unique to the CPI and the categories themselves do not correspond to the North American Industry Classification System (NAICS), other price indexes, or other statistics.

Eight major groups and examples of categories in each follow:

- Food and beverages (breakfast cereal, milk, coffee, chicken, wine, full service meals, snacks)
- Housing (rent of primary residence, owners' equivalent rent, utilities, bedroom furniture)
- Apparel (men's shirts and sweaters, women's dresses, baby clothes, shoes, jewelry)
- Transportation (new vehicles, airline fares, gasoline, motor vehicle insurance)
- Medical care (prescription drugs, medical equipment and supplies, physicians' services, eyeglasses and eye care, hospital services)
- Recreation (televisions, toys, pets and pet products, sports equipment, park and museum admissions)
- Education and communication (college tuition, postage, telephone services, computer software and accessories)
- Other goods and services (tobacco and smoking products, haircuts and other personal services, funeral expenses)

Additionally, for analytical purposes, the CPI is also divided into food, energy, and all items less food and energy. The CPI for all items less food and energy gets considerable attention as a measure of underlying "core" inflation, which is not subject to the volatile movements of food and energy prices. A third structure separates the CPI into

commodities and services, with commodities further divided into durables and nondurables. All three structures are comprehensive, with the subcomponents in each structure aggregating to the all items index.

Geography

The CPI is computed for several geographic areas, which are either large metropolitan areas or groups of smaller metropolitan areas in the same region. CPI data are published for 23 metropolitan areas, 4 geographic regions, and 9 divisions as defined by the Census Bureau. Additionally, the CPI is published for two different size classes (above and below 2.5 million in population) at the national and regional level. Appendix 1, geographic sample, shows a list of published areas.

Population

The CPI is computed officially for two different populations. The CPI for all urban consumers (CPI-U) is the broadest measure and is the most widely used CPI. It is based on the expenditure patterns of a sample of urban consumers representing 93 percent of the population.

Not included in the CPI are the spending patterns of people living in rural nonmetropolitan areas, farm households, people in the Armed Forces, and those in institutions, such as prisons and mental hospitals. Consumer inflation for all urban consumers is measured by two indexes—namely, the CPI-U and the Chained Consumer Price Index for All Urban Consumers (C-CPI-U).

The Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W) is based on the expenditures of urban households included in the CPI-U definition that also meet two additional requirements: more than one-half of the household's income must come from clerical or wage occupations, and at least one of the household's earners must have been employed for at least 37 weeks during the previous 12 months. The CPI-W population represents about 29 percent of the total U.S. population and is a subset of the CPI-U population.

Seasonality

Expenditure patterns can change dramatically during the year, for example, around holidays. A statistical process called seasonal adjustment can be done to remove typical seasonal influences, which helps economists better identify underlying trends. CPI publishes both seasonally adjusted data and not seasonally adjusted data. Seasonally adjusted data are typically used to understand and analyze month-to-month price change; not seasonally adjusted data are typically used for official purposes, including escalation of government payments and for studying longer term price movement. Seasonally adjusted data are only published at the nationwide level, and not all categories are seasonally adjusted.

Base periods

The base period is when the time series of index values is normalized to 100. The current standard base for most categories is 1982–84=100. Some indexes are also published on an alternate base, which is most often 1967=100. Some newer indexes have more recent bases, which are typically the first month of their computation equaling 100.

An index is fully described by its item category, geographic location, population, seasonality, and base. In its databases, BLS uses descriptions with codes for each of these variables to identify a specific CPI. Additional information about Series ID codes is available at the BLS Help & Tutorials page.

Scope and exclusions

The CPI provides an approximation to a conditional cost-of-living index, pricing consumer goods and services. Free goods, the quality of the environment, goods provided by the government at no cost, and the value of leisure time, are all out of scope, despite affecting the cost of living as broadly defined.

Excluded goods and services

The CPI covers the consumption sector of the U.S. economy, which is defined as the purchase of goods and services for use by households. Consequently, the CPI excludes investment items, such as stocks, bonds, real estate, and business expenses. Life insurance is also excluded for this reason, although health, household, and vehicle insurance are in scope. Employer provided in-kind benefits are viewed as part of income rather than consumption. Purchases of houses, antiques, and collectibles are viewed as investment expenditures and therefore excluded. Gambling losses, fines, cash gifts to individuals or charities, and child support and alimony payments also are out of scope. Interest costs and finance charges are also out of scope. The CPI excludes illegal goods and services and the value of home-produced items because of the practical difficulties of collecting the data.

Government-provided and government-subsidized items

The CPI treats any changes to fees that the government charges for items, such as admission to a national park, as in-scope changes in price. The CPI also counts the price of subsidized items that is available to the general public. For example, governments may subsidize local transit operations. If the subsidy is cut and the fare is raised, the CPI will reflect this as a price increase. On the other hand, the CPI does not reflect changes to means-tested subsidies (dependent on the recipient's income), such as the Supplemental Nutrition Assistance Program or Section 8 housing allowances. Changes in such subsidies are treated as changes to the recipient's income and are out of scope.

Taxes

The CPI excludes income tax and other direct taxes; however, it does include the effects of changes in sales taxes and other indirect taxes paid on consumer products. No attempt is made to reflect changes in the quantity or quality of government services paid for through taxes.

Last Modified Date: November 24, 2020

Data Sources

The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by consumers for a representative basket of consumer goods and services. The CPI measures inflation as experienced by consumers in their day-to-day living expenses. The CPI is used to adjust income eligibility levels for government assistance, federal tax brackets, federally mandated cost of living increases, private sector wage and salary increases, and consumer and commercial rent escalations. Consequently, the CPI directly affects hundreds of millions of Americans.

The CPI is created from a series of interrelated surveys. The CPI requires

- a geographic sample, which is a set of areas where prices will be collected;
- a survey of consumer expenditures to create and appropriately weight a market basket of goods and services to be priced and to create a sample of outlets in which prices are collected; and
- samples of prices for commodities, services, and housing.

Geographic sample

Using 2010 census population data, we select the urban areas from which data on prices are collected and choose the housing units within each area that are eligible for use in the shelter component of the CPI. The census data also provide information on the number of consumers represented by each area selected as a CPI price collection area. Additional information on the process of creating the geographic sample is available in the design section.

Consumer expenditure data

The CPI seeks to measure the change in the cost of living by measuring the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services. For the CPI to be accurate, the market basket must correspond to what consumers are actually purchasing, and the different categories of items must be weighted to reflect their proportions in consumers' budgets.

The CPI uses data from the Consumer Expenditure (CE) survey to determine the weights of the different categories of goods and services in the CPI. The CE survey collects data on the out-of-pocket expenses spent to acquire all consumer products and services. The CPI uses the CE data to identify the goods within the CPI's scope. Information about the scope of the CPI is available in the concepts section. Annual CE data are used for the CPI-U and CPI-W weights; these expenditure weights are updated biennially. For example, annual CE data from 2015 and 2016 were used to construct a set of weights that were implemented in the CPI at the end of 2017 and were used through the end of 2019. Additional information about the CE survey is available in the Consumer Expenditure section of the BLS Handbook of Methods.

Price data

The most fundamental data in the CPI are prices. CPI price data are collected via two surveys: one survey collects prices for commodities and services and the other survey collects prices for rent.

Commodities and services survey

The CPI survey collects about 94,000 prices per month to compute indexes for commodities and services. Approximately two-thirds of price collection in the CPI is done by personal visits of CPI data collectors to brick-and-mortar stores. The remaining data are collected by telephone or on the outlet's website. In some cases, these data are supplemented by data provided from other sources.

The outlets where prices are collected are selected based on data from the CE survey. These outlets may be brick-and-mortar stores or websites (e-commerce); currently, about 8 percent of CPI quotes are collected from outlet websites.

Some secondary sources are also used in constructing the CPI sample. For example, data from the U.S. Department of Transportation database are used to construct the sample of fares in the airline fares index.

Housing survey

The CPI survey collects about 8,000 rental housing unit quotes each month to compute the indexes for the housing component. It uses price quotes for rent and homeowners' equivalent rent (an estimate of the implicit rent that owner occupants would have to pay if they were renting their homes) to compute estimates of price change. Because rents change rather infrequently, the CPI program collects rent data from each sampled unit every 6 months. Collecting rent data less frequently allows for a much larger sample. The CPI divides each area's rent sample into six sub-samples called panels. The rents for panel 1 are collected in January and July; panel 2, in February and August, etc. Rents are collected by personal visit or phone.[1]

Repricing and quality adjustment

Prices for each item in the commodities and services survey are collected either every month or every other month, depending on the type of good or service and its location. Food at home, energy, and selected other items are priced monthly. So, too, are prices for all other commodity and service items in the three largest publication areas: New York, Los Angeles, and Chicago. Elsewhere, prices are collected bimonthly for the remaining commodity and service items; those are assigned to either even- or odd-numbered months for pricing.

Most repricing is done by personal visit from a CPI data collector, but in other cases repricing is done by a website visit or by telephone. If the selected item is available, a data collector records its price and the recorded information is reviewed by commodity analysts who have detailed knowledge about the particular good or service. Unusual price movements are reviewed carefully and checked for validity. The price index formula cannot handle a price of zero (or free), therefore, a zero price is adjusted to a very small price.

If the selected item is no longer available, or if there have been changes in the quality or quantity (for example, a container of orange juice containing 59 ounces instead of 64 ounces) of the good or service since the last time prices were collected, the data collector selects a new item similar to the old item. This is referred to as a substitution.

When substitution occurs, the commodity analyst reviews the new item and price. The new price may be quality adjusted for use in index computation. Conceptually, the CPI seeks to be a constant-quality measure, though accurately quantifying quality change may not always be possible. Detailed information about quality adjustment procedures is in the calculation section.

Alternative data sources

Although most of the prices used to compute the CPI are collected by BLS through the process decribed above, in some cases these data are supplemented by data from other sources.

Airline fares

Data from the U.S. Department of Transportation database are used to construct the sample of fares in the airline fares index.

Apparel and household goods

Among the many firms that participate in the CPI survey, one firm provides BLS with a large volume of price data rather than allowing data collectors to collect data in stores. Additional information on this methodology is available in the paper " *Big Data in the U.S. Consumer Price Index: Experiences & Plans*."

Postage

For sample selection, the delivery of household mail by type of postal service and postal zone is determined by the United States Postal Service (USPS) Household Diary Survey. We collect monthly prices from the price list on the USPS public website.

Prescription drugs

One firm provides BLS with a large volume of price data rather than allowing data collectors to collect data in stores. Additional information on how BLS prices prescription drugs is available in the medical care factsheet.

Used cars and trucks

For used cars and trucks, both the sample and prices come from alternative sources. The current CPI sample of used cars and trucks comes from the J. D. Power Information Network, which is a network of car dealers who report sales of used vehicles to the J. D. Power Company. From the universe of 2- through 8-year-old vehicles, we choose a sample of 480 vehicles. The 480 observations are replicated in all of the CPI areas (after tax adjustments). The sample is updated by one model year each September, October, or November to maintain the same-age vehicles over time. If a production model is discontinued, it is replaced by a comparable model and a complete resampling is conducted every 5 years.

All price information for used cars and trucks in the CPI comes from the National Automobile Dealers Association Official Used Car Guide (All prices are adjusted for depreciation of the vehicle. Additional information about how BLS prices used cars is available in the used cars and trucks factsheet.)

Confidentiality

Data from the pricing surveys are collected under pledges of confidentiality, and BLS is bound by law to protect the confidentiality of respondents. Data collection and security procedures are governed by provisions of the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA).

The BLS confidentiality pledge

Respondents to BLS surveys will receive the confidentiality pledge assuring them of BLS commitment to keep their information secure. For the price surveys (both the commodities and services survey and the housing survey), the pledge is a follows:

The Bureau of Labor Statistics, its employees, agents, and partner statistical agencies, will use the information you provide for statistical purposes only and will hold the information in confidence to the full extent permitted by law. In accordance with the Confidential Information Protection and Statistical Efficiency Act of 2002 (Title 5 of Public Law 107-347) and other applicable Federal laws, your responses will not be disclosed in identifiable form without your informed consent.

Multiple confidentiality issues arise in the production of the CPI. One set of issues arises out of the need to prevent unauthorized access of data that are embargoed, or yet to be released to the public. Because the CPI data can affect financial markets, it is essential to ensure that no one without authorization has access to the data before release. BLS personnel who do see the data ahead of time are restricted by law from engaging in certain financial transactions during the period where they have seen data that are not public. Pre-release data are encrypted and always kept on secure servers, with any hard copies locked in secured areas.

In some cases, publication of indexes could risk revealing the pricing behavior of a respondent and publication of that data are suppressed according to specific thresholds.

Another set of issues relates to the protection of respondent identifying information (RII) and personally identifiable information (PII). We do not release price data from the sample to the public, nor do we confirm or deny that any specific product or brand is in the price sample, or that any specific store or seller is in the outlet sample.

CPI employees are trained annually on procedures to protect the security of embargoed data and the privacy of respondents. Additional information about CPI confidentiality procedures is available at the BLS Confidentiality of Data Collected by BLS for Statistical Purposes page.

NOTES

[1] For more information on *Owners' equivalent rent of primary residence (OER)* and *Rent of primary residence (Rent)*, please see https://www.**bls.gov/cpi**/factsheets/owners-equivalent-rent-and-rent.pdf." *Monthly Labor Review,* April 2014

Last Modified Date: November 24, 2020

Design

The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by consumers for a representative basket of consumer goods and services. The CPI measures inflation as experienced by consumers in their day-to-day living expenses. The CPI is used to adjust income eligibility levels for government assistance, federal tax brackets, federally mandated cost of living increases, private sector wage and salary increases, and consumer and commercial rent escalations. Consequently, the CPI directly affects hundreds of millions of Americans.

Target Population

The CPI measures the average price change over time for a market basket of goods and services for two target populations: All Urban Consumers (CPI-U population) and Urban Wage Earners and Clerical Workers (CPI-W population).

Both the CPI-U and the Chained CPI (C-CPI-U) use the CPI-U population. The CPI-U population constitutes about 93 percent of the U.S. population, and covers households in all areas of the United States, specifically, all urban households in Core-Based Statistical Areas (CBSAs) and in urban places of 10,000 inhabitants or more.[1] Not covered are people living in rural nonmetropolitan areas, in farm households, on military installations, in religious communities, and in institutions such as prisons and mental hospitals.

The CPI-W population is a subset of the CPI-U population. The CPI-W consists of all CPI-U population households in which at least one of the members has been employed for 37 weeks or more in an eligible occupation and for which 50 percent or more of the household income must come from wage earnings associated with an eligible occupation. Eligible occupations include clerical workers, sales workers, protective and other service workers, laborers, and construction workers. The CPI-W population excludes households of professional and salaried workers, part-time workers, the self-employed, and the unemployed, along with households with no one in the labor force, such as those of retirees. The CPI-W share of the total U.S. population has diminished over the years and is now about 29 percent of the total U.S. population.

The sample

The CPI collects prices for approximately 80,000 goods and services. Prices are collected each month in 75 urban areas across the country from about 6,000 housing units and approximately 23,000 retail establishments—department stores, supermarkets, hospitals, gas stations, and other types of stores and service establishments. All taxes directly associated with the purchase and use of items are included in the index. Prices of fuels and a few other items are obtained every month in all 75 locations, while prices of most other commodities and services are collected every month in the three largest geographic areas (Chicago, Los Angeles, and New York) and every other month in other areas. Prices of goods and services are obtained primarily through personal visits or telephone calls by BLS data collectors, though some prices are collected directly from websites. Definitions of entry level items are available in appendix 2.

In calculating the index, price changes for the various items in each location are averaged together using weights that represent their importance in the spending of the appropriate population group. Local data are then aggregated to obtain a U.S. city average. For the CPI-U and CPI-W, separate indexes are also published by size of city, by census region, by census division, for cross-classifications of regions and population-size classes, and for 23 local areas. For the C-CPI-U, data are published only at the national level. The CPI-U and CPI-W indexes are considered final when released, but the C-CPI-U index is issued in preliminary form and is subject to three quarterly revisions before the final version is published 10-12 months after the initial release.

In order to select stores and items to calculate indexes, BLS implements a sampling design that constructs the sampling frames from which a random sample of stores, consumer items, and expenditure weights can be drawn. This section describes the basic elements of CPI sampling design and the steps taken to implement the design.

Multi-stage sampling design: areas, outlets, and items

The CPI sample-design process involves multiple stages. In the first stage, a sample of geographic areas is selected. In subsequent stages, BLS selects a sample of outlets in which area residents make retail purchases, a sample of specific retail goods and services that area residents buy, and a sample of residential housing units. The samples are rotated on a regular basis; the geographic sample has traditionally been rotated once after each decennial census.

Area sample

Effective with the 2018 redesign based on the 2010 census, the current geographic sample (appendix 1) was introduced over a multi-year span beginning in 2018. The area definitions are based on the 2013 Office of Management and Budget's (OMB) CBSAs.[2]

Area sampling steps

1. Determine sample classification variables.

In the current sample design, areas were first classified into one of nine census divisions: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific. The census divisions represent a further breakdown of census regions. Each area was also classified into one of two population-size classes: self-representing or non-self-representing. Areas above 2.5 million are defined as self-representing; their weight in the CPI corresponds to their population relative to the U.S. population. Areas below 2.5 million represent not only themselves, but other areas in their region and size class.

Exhibit 1. Census Regions and Divisions of the United States

2. Construct primary sampling units.

The current definitions assign counties surrounding an urban core area to geographic entities (the CBSAs). The assignment is based on each county's degree of economic and social integration (as measured by commuting patterns) to the urban core. There are two types of CBSAs: metropolitan and micropolitan. A metropolitan CBSA has an urban core of more than 50,000 people, and a micropolitan CBSA has an urban core of 10,000 to 50,000 people. Note that CBSAs may cross state borders.

Metropolitan and micropolitan CBSA definitions were used for non-self-representing areas. With the introduction of the CBSA concept to the CPI, the CPI-U covers 93 percent of the U.S. population reflected in the 2010 census. The area sample frame comprises 381 metropolitan CBSAs, representing approximately 85 percent of the population, and 536 micropolitan CBSAs, representing approximately 9 percent of the population.

3. Determine the number of sampled PSUs.

The CPI program maintains 75 primary sampling units (PSUs). These PSUs include 21 areas whose population is greater than 2.5 million and 2 additional areas: Anchorage, AK, and Honolulu, HI. Anchorage represents all CBSAs in Alaska, and Honolulu represents all CBSAs in Hawaii. These CBSAs are unique because the locations of both states make price change in their markets geographically isolated from that in other markets, so the CBSAs in Alaska and Hawaii are treated as separate geographic strata. These 23 self-representing PSUs are combined with 52 non-self-representing PSUs to form 75 total PSUs. For purposes of index calculation, the 75 PSUs are consolidated into 32 index areas. Thus, the current area design yields 7,776 basic indexes (32 index areas by 243 item strata) for the U.S. all-items CPI.

4. Determine stratification variables.

To best create a sample of areas that represent the entire population, the CPI uses a stratified sample. Many demographic stratification variables were considered, and models including different variables were investigated. The final model selected uses four variables: longitude, latitude, median property value, and median household income.

5. Allocate sample and assign PSUs to strata.

The 23 self-representing PSUs account for approximately 39 percent of the total U.S. population and about 42 percent of the CPI-U population. There are 52 non-self-representing PSUs, which represent the remaining 58 percent of the CPI-U population and include both metropolitan and micropolitan areas.

The next phase of the selection process was to assign the non-self-representing PSUs within each census division to strata based on a model of the four stratification variables. The primary objective of the PSU stratification was to minimize the between-PSU component of variance by making the PSUs within each stratum as homogeneous as possible with respect to the four stratification variables. In addition, to further minimize the variance, strata within each census division had to be kept with approximately the same population.

6. Select a sample of PSUs.

The final step of the selection process was to select one PSU per stratum. Before making that final selection, we chose to employ a sample-overlap procedure, which is intended to increase the expected number of non-self-representing areas reselected in the new design. Additionally, we implemented a controlled-selection procedure, which aims to conduct a random sample in a way that increases the probability of selecting certain preferred combinations of PSUs. After adjusting the sample selection probabilities with the use of the Ernst sample-overlap procedure and employing controlled selection for the micropolitan areas, we randomly selected one PSU per stratum.[3]

7. New area design implementation plan.

After selecting the final area design, BLS determined the process for implementing the new geographic sample into the three surveys used to construct the CPI. The surveys are the Consumer Expenditure (CE) survey, the Commodity and Services survey, and the Housing Survey.

For the 2018 area revision, the CE fully converted to the new sample in 2015. However, for the other surveys, which are directly managed by the CPI program, the 21 new PSUs have been divided into groups and the new PSUs will be introduced over a multiple-year span. This rotation process distributes the cost of introducing new PSUs into the Housing and Commodities and Services surveys, avoiding a spike in data collection costs before the full conversion to the new area design.

The calculation of price indexes under the new area design began in January 2018, with the introduction of the first set of new PSUs in the sample. Existing PSUs scheduled to be rotated out of the sample later in the implementation process will be used as proxy candidates for new PSUs rotating in late, until the complete set of new PSUs has been rotated into the sample. An ideal proxy for a new PSU was considered to be one of the PSUs to be dropped from the sample within a new PSU geographic stratum. If PSUs that had been dropped were not available, a proxy was identified through nearest-neighbor rules, with the constraint that the proxy falls within 200 miles of the new PSU. If no eligible proxy existed, the new PSU was considered to be a “geographic hole” within the new area structure. There were eight new PSUs with no eligible proxy that were given priority in the rotation schedule.

Outlet sample

The outlet sample for most items in the CPI is developed with data from the CE Survey. The survey furnishes data on retail outlets from which metropolitan and micropolitan households purchased well-defined groups of commodities and services to be priced in the CPI.

Commodities and services are grouped into sampling categories based on entry-level items entry-level items as defined in the CPI classification structure. Some categories consist of only one entry-level item, while others consist of more. Entry-level items are combined into a single category when the commodities or services generally are sold in the same outlets; for example, boys' outerwear and boys' shirts and sweaters are both in the same category.

Additional information is available in the CE Survey’s section of the BLS Handbook of Methods.

Procedures for selecting items within outlets

Each outlet is assigned a number of entry-level items for price collection. A data collector visits each selected outlet and uses a multistage probability selection technique to select specific items from among all the items the outlet sells that fall within the entry-level item definitions. Additional information on categories and entry-level item titles is available in the entry-level item definitions spreadsheet (appendix 2, ELI list), the CE categories spreadsheet (appendix 3, CE categories), the non-CE categories spreadsheet (appendix 4, NonCE categories), and the CE_CPI concordance spreadsheet (appendix 5, CE CPI concordance).

Data collectors first identify all of the items included in an entry-level item definition and all of the items that are offered for sale by the outlet. When there are a large number of items in the entry-level items, the data collector groups them by common characteristics, such as brand, size, or type of packaging. With the assistance of the respondent for the outlet, the data collector assigns probabilities of selection to each group.

The probabilities of selection are proportional to the sales of the items included in each group. The data collector may use any of the following procedures to determine the proportion of sales:

- Percents: The percent that a specific group represents of the total dollar sales of all the groups listed in a specific stage of disaggregation
- Ranks: An ordering by the respondent of the groups from largest to smallest in terms of dollar sales
- Dollar volume seller: Method used for the respondent to identify the largest dollar volume selling unique item from the previous group selected in disaggregation
- Ranked selling space: eligible items are ranked by the amount of space they occupy within the store
- Equal probability: Assigns percentages to each of the groups listed in a dissagregation step solely on the basis of the number of groups

After assigning probabilities of selection, data collectors use a procedure to randomly select one group. They then identify all items included in the selected group, form groups of those items based on the characteristics they have in common, assign probabilities to each group, and randomly select one. Data collectors repeat this process through successive stages until reaching a unique item and describe the selected unique item on a checklist for the entry-level item. Checklists contain the descriptive characteristics necessary to identify the item among all items defined within the entry-level item.

These selection procedures ensure that there is an objective and efficient probability sampling of CPI items other than shelter. They also allow broad definitions of entry-level items, so that the same unique item 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 entry-level items is a significantly higher likelihood of finding a priceable item within the definition of the entry-level item in the sample outlet.

The selection process is completed during the visit to the outlet to obtain the price for the selected item. Subsequently, personal visits, telephone calls, or website visits are made, either monthly or bimonthly, to make sure that the item is still sold and to obtain its current price.

Shelter

The CPI Housing Survey provides the data needed to measure price change for the two housing component indexes: owners' equivalent rent of primary residence (OER) and rent of primary residence (Rent). The Housing Survey follows the rents of a sample of renter-occupied housing units selected to represent both renter- and owner-occupied housing units in the urban United States.

We continuously update the sample of rented housing units by replacing one-sixth of the rented housing unit sample every year on the basis of the latest available U.S. Census Bureau data.

Collecting a large sample less frequently is more efficient for the Housing Survey because rent prices are not as volatile as most other consumer prices. This efficiency is accomplished by assigning each selected neighborhood (called a segment) in a pricing area to one of six panels, each of which represents a subsample of each pricing

area and provides sufficient information for the monthly owners' equivalent rent of primary residence and rent indexes. Each month, a panel is priced, with all six panels priced twice a year: panel 1 is collected in January and July, and panel 2 is collected in February and August, and so on. Every month, we collect rent prices and other information for one panel and the 6-month price ratio is computed (the current rent divided by the rent 6 months ago) for each unit in the panel. The measures of price change for the two housing components are based on weighted averages of these rent ratios.

The 2018 geographic sample of the CPI partitioned the urban United States into 32 CPI areas and selected 75 pricing areas. These areas were metropolitan and micropolitan areas and were selected using probability proportional to size (PPS) sampling, which was the size of the 2010 population. For the Housing Survey, CPI pricing areas are further partitioned into neighborhoods called segments, formed from one (in most cases) or more U.S. Census Bureau block groups and containing at least 50 housing units in large (population size greater than 2,500,000) self-representing pricing areas and at least 30 in smaller non-self-representing pricing areas. With the use of PPS, a sample of segments was selected in each area, in which the size measure was the sum of renters' actual rents and owners' estimated implicit rents. The Census Bureau provides the number of renters, the average rents, and the number of owners by block group, whereas BLS estimates the average owners' implicit rents. An average of about five rental housing units is selected within each segment.

Housing sample

In 2010, the CPI undertook a three-stage effort to improve the Housing Survey. The first and second stages used the 2000 census. The third stage is a regular replacement commencing in 2016 and ending in 2022. It will replace the 2000 census-based sample with one based on the American Community Survey using 2010 Decennial census geography. This stage will continue into the future and—for the first time—the CPI Housing Survey will have a process that keeps its sample continuously updated.

BLS staff use purchased address lists and a mail prescreening survey to locate housing units in the segments. The lists indicate the probability that an address is owner-occupied and the addresses provide a means of determining whether an address is a commercial establishment. This information isused to determine sampling rates for the mail prescreening survey and to determine if selected addresses are commercial or residential and, if they are residential, their tenure (owner- or renter-occupied). Only addresses the survey identifies as renter-occupied as well as those with no response are sent out for data collectors to screen.

Staff must find an eligible respondent for each address. During screening, the interviewing software directs the data collectors through a structured series of questions to verify that the unit is renter-occupied. Data collectors further determine the following:

- the unit is the primary residence of the occupant,
- the occupant is not a relative of the landlord,
- the unit is not institutional,
- the unit is not public housing, and
- the unit is not an assisted-living facility with activities of daily living provided to an occupant.

These questions help the data collectors determine if a selected address is eligible for the Housing Survey sample.

The data collector has a multi-month period to screen and initiate the units in a segment; units that they are not able to screen (usually because the field agents fail to contact a respondent) go back out “on panel” for another screening attempt. This process should yield an expected number (usually five) of in-scope housing units in each segment that will be initiated into the Housing Survey sample.

Initiation and pricing

Once a selected address has been successfully screened, the data collector immediately proceeds to initiate the housing unit into the Housing Survey sample. Initiation is the initial collection of rent data, which the data collector obtains by asking another structured series of questions. These data include the rent that is paid and specific housing services that are associated with the unit and are the basis for all calculations of rent change that will occur during the life of the housing unit in the housing sample.

Once a unit is initiated, it will be priced on panel every 6 months. In addition, any in-scope units that are not successfully initiated go back on panel for another attempted initiation. A housing unit’s initiation generally does not occur in its on-panel month, so the housing unit must be priced on panel for two cycles to provide a 6-month interval before price changes can be used in the indexes. During initiation and during each pricing, BLS collects

- contract rent and rental period (monthly, bimonthly, weekly, or for a specified number of days);
- utilities, facilities, and any other such items included in the rent;
- any subsidies (e.g., Section 8) or reductions in the rent in exchange for services the tenant provided;
- any extra charges included in the contract rent for optional items, such as parking;
- the number of rooms, type of housing structure, and other physical characteristics; and
- equipment used for air conditioning and fuels used for heat and hot water.

In addition, to ensure that the unit remains in scope, we ask the screening questions every 2 years or when a change of occupant occurs.

NOTES

[1] Core Based Statistical Areas (CBSAs) consist of the county or counties or equivalent entities associated with at least one core (urbanized area or urban cluster) of at least 10,000 population, plus adjacent counties having a high degree of social and economic integration with the core as measured through commuting ties with the counties associated with the core.

[2] *Revised delineations of metropolitan statistical areas, micropolitan statistical areas, and combined statistical areas, and guidance on uses of the delineations of these areas*, Bulletin No. 13-01 (Office of Management and Budget, February 28, 2013), https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/bulletins/2013/b13-01.pdf.

[3] See https://www.bls.gov/osmr/research-papers/1999/pdf/st990060.pdf.

Last Modified Date: November 24, 2020

Calculation

The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by consumers for a representative basket of consumer goods and services. The CPI measures inflation as experienced by consumers in their day-to-day living expenses. The CPI is used to adjust income eligibility levels for government assistance, federal tax brackets, federally mandated cost of living increases, private sector wage and salary increases, and consumer and commercial rent escalations. Consequently, the CPI directly affects hundreds of millions of Americans.

Estimation of price change in the Consumer Price Index

In the Consumer Price Index (CPI), the urban areas of the United States are divided into 32 geographic areas, called index areas. The set of all goods and services purchased by consumers is divided into 211 categories called item strata: 209 Commodities and Services item strata, plus 2 housing item strata. The number of basic items used for the calculation of aggregate indexes is larger than this, at 243, because the entry level item (ELI) level is used for the calculation of basic cells for health insurance retained earnings (item code SEME) rather than the higher item stratum level. This results in 7,776 (32 x 243) item-area combinations.

Estimation of price change in the CPI is done in two stages. The first stage of CPI estimation is to calculate basic indexes for each of the 7,776 item-area combinations that compose the CPI. For example, the CPI-U series for electricity in the Philadelphia-Camden-Wilmington, PA-NJ-DE-MD CPI area is a basic index. The weights for this first stage come from the sampling frame for the item strata in the index area. Then, at the second stage, we calculate aggregate indexes by averaging across subsets of item-area combinations. For example, the all items index for Philadelphia is the aggregate of all 243 basic index series in that index area. Similarly, the U.S. city average index for electricity is the aggregate of the basic indexes for electricity in each of the 32 index areas. The U.S. city average all items CPI is the aggregate of all basic indexes. For the CPI-U and CPI-W, the weights for the second stage of aggregation are the biennial reference-period expenditures on the item strata in the index area, as calculated using expenditure data from the Consumer Expenditure Surveys.

Estimation of price change for commodities and services other than shelter

For the majority of the 209 C&S strata, most information on price change comes from the commodities and services pricing survey. A few C&S strata including those for airline fares, intercity train fares, and used vehicles, use secondary sources of data on prices for their samples. For 24 strata with small weights, price movements are imputed from related strata.

Price relatives

Each month, the processing of the C&S survey data yields a set of price relatives, which are measures of short-term price change for all basic indexes. The CPI uses an index number formula to obtain an average price change for the items in each basic index's sample. Most item strata use the geometric mean index formula, which is a weighted geometric mean of price ratios (the item's current price divided by its price in the previous period) with weights equal to expenditures on the items in their sampling periods.

Calculations for a limited number of strata use a modified Laspeyres index number formula, which is a ratio of a weighted arithmetic mean of prices in the current period to the same average of the same items' prices in the previous period, with estimated quantities of the items purchased in the sampling period serving as weights. The following strata use the Laspeyres formula:

- Selected shelter services (housing at school, excluding board)
- Selected utilities and government fees (electricity, residential water and sewerage maintenance, utility (piped) gas service, state motor vehicle registration and license fees)
- Selected medical care services (prescription drugs, physicians' services, hospital services, dental services, services by other medical professionals, and nursing homes and adult day care)

Each month, the estimation system uses the following formulas to compute price relatives for each item-area combination *(i,a)*.

The price relative (using a geometric mean formula) is given by

$$_{(i,a)}P^{G}_{t,t-1} = \prod_{j \in (i,a)} \left(\frac{P_{j,t}}{P_{j,t-1}}\right)^{w_{j,b}}$$

The price relative (using a Laspeyres formula) is given by

$$_{(i,a)}P^{L}_{t,t-1} = \frac{\sum_{j \in (i,a)} P_{j,t}\, w_{j,b}}{\sum_{j \in (i,a)} P_{j,b}\, w_{j,b}}$$

where,

$_{(i,a)}P^{G}_{t,t-1}$ = the geometric price relative for the item-area combination *(i,a)* from the previous period *t–1* to the current period *t;*

$_{(i,a)}P^{L}_{t,t-1}$ = the Laspeyres price relative for the item-area combination *(i,a)* from the previous period *t–1* to the current period *t;*

$P_{j,t}$ = the price of item *j*, which is a member of item stratum *i*, for which a price quote is being collected in area *a*, observed in period *t*;

$P_{j,t-1}$ = the price of the same item *j* in period *t–1*;

$P_{j,b}$ = an estimate of item *j'* s price in the base period; and

$w_{j,b}$ = item *j's* weight in the base period.

The product and sums in the formulas presented above are taken over all price quotes which are usable for estimation in the item-area combination *(i,a)*. It is important that the price of each quote be collected (or estimated) in both periods in order to measure price change.

Quote weights

For each individual quote, the weight, or each quote's share of the average daily expenditure on the ELI in the primary sampling unit (PSU), is given by $w_{j,b,}$ which is computed as

$$w_{j,b,} = \frac{AEfg\eta}{N}$$

where

A = the proportion of CE expenditures for the ELI relative to the entire item category within the Census region;

E= estimate of the total daily expenditure for the item category in the PSU by people in the CPI-U population (called the basic weight);

f = a duplication factor that accounts for any special subsampling of outlets and quotes;

g = a geographic factor used to account for differences in the index area's coverage when the CPI is changing its area design;

N = the number of quotes planned for collection in the item stratum PSU, which is also the sum of duplication factors for all sampled quotes in the item stratum PSU;

B = the proportion of CE expenditures for the ELI relative to the item stratum within the region; and

η = a nonresponse adjustment factor calculated as the quantity $1+\frac{y}{n-y}$ where y is the sum of duplication factors for uninitiated quotes and n is the number of quotes in the sample design in the ELI-PSU. This is the ratio of planned quotes to quotes with usable prices in both period t and period $t–1$ for the ELI-PSU.

Base-period prices

In the modified Laspeyres formula used for C&S items, the quote weight is divided by an estimate of the item's price in the sampling period to obtain an estimated quantity. An item's base period occurs sometime before its outlet's initiation, so one cannot observe its base-period price directly. Instead, the price is estimated from the item's price at the time the sample was initiated and the best available estimates of price change for the period from the base period to the initiation period.

The price of an item, j, in the base period is given by

$$p_{j,0} = \frac{p_{j,1}}{r_{j,1}}$$

where

$p_{j,0}$= the price of item j at the time of initiation (period 0),

$IX_{j,o}$ = the value of the price index most appropriate for item *j* at the time of initiation, and

$IX_{j,b}$ = the value of the same price index in the base period.

Item replacement and quality adjustment

One of the more difficult 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 a measurement through time of the cost of purchasing an unchanging, constant-quality set of goods and services. In reality, products disappear, products are replaced with new versions, and new products emerge.

When a data collector finds that he or she can no longer obtain a price for an item in the CPI sample (often because the outlet permanently stops selling it), the data collector uses the CPI item replacement procedure to find a new item. Each priced item stratum in the CPI contains one or more ELIs. CPI commodity analysts have developed checklists that define further subdivisions of each ELI. When seeking a replacement in a retail outlet, the data collector first uses the checklist for the ELI to find the item sold by the outlet that is the closest to the previously priced item. Then the data collector describes the replacement item on the checklist, capturing its important specifications. The CA assigned to the ELI reviews all replacements and selects one of three methods to adjust for quality change and to account for the change in item specifications.

The following example describes the most common type of quality adjustment problem. Assume that a data collector in period *t* tries to collect the price for item *j* in its assigned outlet and is not able to do so because the outlet no longer sells this item. (A price for item *j* was collected in period *t–1*.) The data collector then finds a replacement item and collects a price for it. This replacement item becomes the new version *v+1* of item *j*. The commodity analyst decides how the CPI treats the replacement. The commodity analyst has the descriptions of the two versions of item *j*. In addition, he or she has the *t–1* price, $P^{v}_{j,t-1}$, for the earlier version *v* and the period *t* price, $P^{v+1}_{j,t}$, of the replacement version *v+1*. The following matrix displays the information available to the commodity analyst:

Version	**Period *t–1* price**	**Period *t* price**
Old version *v*	$P^{v}_{j,t-1}$	–
Replacement version *v+1*	–	$P^{v+1}_{j,t}$

To use the item in index calculation for period *t*, it is necessary to have an estimate of $P^{v}_{j,t}$, which is the price of the earlier version *v* in the current period *t*. If there is no accepted way of estimating either $P^{v+1}_{j,t-1}$ or $P^{v}_{j,t}$, the observation for item *j* is left out of the index calculation for period *t*, meaning that the observation is treated as a nonresponse handled by imputation.

The three methods from which a commodity analyst can choose to handle the replacement follow.

Direct comparison

If the original and replacement items are essentially the same, the CA deems them directly comparable, and the price comparison between the items is used in the index. In this case, it is assumed that no quality difference exists between the versions.

Direct quality adjustment

The most explicit method for dealing with a replacement item with a difference in quality is to estimate the value of the differences. The estimate of this value is called a quality adjustment amount, $QA_{j,t-1}$. In this case,

$$QA_{j,t-1} = P_{j,t}^{v+1} - P_{j,t-1}^{v}$$

where

$P_{j,t}^{v+1}$ = the period *t* price of the replacement version *v+1,* and

$P_{j,t-1}^{v}$ = the period *t–1* price for the earlier version *v*.

Sources of direct quality adjustment information include observable factors such as size or weight, manufacturers' cost data, and hedonic regression models.

Imputation

Imputation is a procedure for handling missing information. The CPI uses imputation for a number of cases, including respondent refusals, items which are out of season or unavailable for some other reason, and the inability to make a satisfactory estimate of the quality change. Replacement items that can be neither directly compared nor quality adjusted are called noncomparable. For noncomparable replacements, an estimate of constant-quality price change is made by imputation. There are two imputation methods used in the CPI: cell-relative imputation and class-mean imputation.

Cell-relative imputation

If there is no reason to believe that the price change for an item is different from the price change observed for the other items in its basic index, the cell-relative method is used to impute the change. This method is used for missing values, because no information is available about the observation in such cases. For noncomparable substitutions, this method is common for food and service items. The price change between the original item and the noncomparable replacement item is assumed to be the same as the average price change of all similar items in 1 month for the same geographic area, (i.e., the same as the average price change for the basic cell for that ELI and PSU). When there is a new version of the item that is not comparable to the previous version, a price of the new version is available. That price is not used in calculations for period *t*, but will be used in the subsequent period *t+1* as the previous price.

Class-mean imputation

Some C&S item strata use a class-mean imputation for many noncomparable replacements, primarily in the item strata for vehicles, for other durables, and for apparel. The logic behind the class-mean procedure is that price change is closely associated with the annual or periodic introduction of new lines or models for many items. For example, at the introduction of new model-year vehicles, there are often price increases while, later in the model year, price decreases are common. The CPI uses the quality adjustment method as frequently as possible to handle item replacements that occur when product lines are updated. Class-mean imputation is employed in the remaining replacement situations. In those cases, the CPI estimates price change from the price changes of other observations that are going through an item replacement at the same time and that were either quality adjusted directly or judged to be directly comparable. For class-mean imputation, the CPI estimates $P_{j,t}^{v}$, which is an estimate of the current period *t* price for the old version *v*, and uses this estimated current price in the calculation of the price relative for period *t*.

The estimated current-period price is the previous period *t–1* price of the old version multiplied by a specially constructed price relative for the class cR:

$$P_{j,t}^{v} = P_{j,t-1}^{v} \cdot cR_{t,t-1}$$

where $cR_{t,t-1}$ is computed with either the geometric mean or Laspeyres formula over the subset of the observations in the ELI of which item *j* is a member. The subset is the class of interest, that is all the comparable and quality-adjusted replacement observations in the same ELI and PSU.

Review and treatment of outlier price changes

All outlier price changes are reviewed by CAs. Outlier price changes, if accurate, are generally included in the calculation of price relatives. Extreme price changes are given upper and lower bounds, say 10 and 0.1.

Estimation of price change for shelter

The rent of primary residence (Rent) index and owners' equivalent rent of primary residence (OER) index measure the change in the cost of shelter for renters and homeowners, respectively. Price change data for these two item strata come from the CPI housing survey. Each month, BLS data collectors gather information from renter units on the rent for the current month and on what services are provided. Rent and OER are each subject to their own unique estimation procedures.

Rent

The rent estimates used in the CPI are contract rents. They are the payment for all services provided by the landlord to the tenant in exchange for rent. For example, if the landlord provides electricity or other utilities, these would be part of the contract rent. The CPI item expenditure weights also include the full contract rent payment. Rents are calculated as the amounts the tenants pay their landlords, plus any rent reductions tenants receive for performing services on behalf of the landlord, plus any subsidy payment paid to the landlord. Reductions for any other reason are not considered part of the rent.

Owners' equivalent rent of primary residence (OER)

The OER approach to price change for owner-occupied housing is designed to measure the change in the rental value of the owner-occupied housing unit; the investment portion is excluded. In essence, OER measures the change in the amount a homeowner would pay in rent or earn from renting his or her home in a competitive market. It is a measure of the change in the price of the shelter service provided to the homeowner by the owner-occupied housing unit.

Unit-level weighting

The housing sample is made up of renter-occupied units from the 2010 Decennial Census of Population and Housing in which higher rent levels (expenditures) have a higher probability of selection. The U.S. Census Bureau provided the numbers of renters and owners and the average rent of renter units in the block groups, and BLS estimated the average implicit rent of the owner units in the block groups. From these pieces of information, CPI calculated the total cost of rent in the block groups from the renter costs and the owner costs in the block groups.

The CPI breaks up each of the 75 CPI PSUs into small geographic areas, which are called segments. Segments are formed from one (in most cases) or more census block groups. The segments are sorted by PSU, state, county, average rent (or rent level) and tract. Blocks are portions of block groups, while tracts are portions of counties, and counties are portions of states. There can be more than one state in a PSU. The census data needed for sample selection are only available at the block group level.

Each calculation begins with a segment weight based on the probability of selecting the segment. A segment weight is the inverse of the probability of selection, where the probability of selection is the total housing cost for the segment multiplied by the number of segments to be selected in the PSU divided by the total housing cost for the PSU.

$$W_s = \frac{n_{PSU} \cdot TC_s}{TC_{PSU}}$$

where

W_s = weight of segment *s*,

n_{PSU} = the number of segments in the PSU,

S = the segment, and

TC_s = the total cost of rents in all block groups in segment *S.*

$$TC_s = \sum_{bg \in s} TC_{bg}$$

where

TC_S = the total cost of rents in all block groups in segment *S*,

S = the segment, and

BG = the block group.

To derive the renter weight in the segment, the segment weight is multiplied by the number of renters in the segment and divided by the number of renters sampled in the segment:

$$RW_s = W_s \frac{R_s}{n_s}$$

where

RW_s = renter weight in segment *s,*

W_s = segment weight,

R_s = the number of renters in segment *s*, and

n_s = the number of renters sampled in segment *s*.

Similarly, the owners' equivalent rents weight is derived by multiplying the segment weight by the number of owners in the segment, and dividing by the number of renters sampled in the segment. Since the housing survey collects rents and not the implicit rents of owners, the ratio of average implicit rent to average rent in the segment is also included in the OER rent weight:

$$OW_s = W_s \frac{O_s}{n_s} \frac{IR_s}{RR_s}$$

where

OW_s = owners' equivalent rents weight in segment *s,*

W_s = segment weight in segment *s,*

O_s = number of owners in segment *s,*

n_s = the number of renters sampled in segment *s,*

IR_s = average implicit rent in segment *s,* and

RR_s = average rent in segment *s.*

Six-month chained estimator

For the rent index, the current month's index is derived by applying the sixth root of the 6-month rent change to the index for the previous month. For the OER index, the current month's index is derived by applying the sixth root of the 6-month OER change to the index for the previous month.

The rent estimator uses the change in the economic rent, which is the contract rent adjusted for any changes in the quality of the housing unit, to estimate the change in the average rent. Due to the panel structure used in the housing sample, the 6-month change in rent is based on sampled, renter-occupied units that have usable 6-month rent changes. The sum of the current period economic rents for each usable unit within a segment, weighted by the renter weight for that segment, is divided by the sum of the weighted economic rents 6 months earlier *t–6*. This ratio is used to represent the 6-month change in rent for all renter-occupied units within a segment.

$$RER^{RENT}_{t-6,t,s} = \frac{\sum_{i \in s} ER_{i,t} \cdot RW_s}{\sum_{i \in s} ER_{i,t-6} \cdot RW_s}$$

where

$RER^{RENT}_{t-6,t,s}$ = relative rent price change between months *t* and *t–6* in segment *s,*

t = current period,

t–6 = period 6 months before the current period *t*,

RW_s = renter weight for segment *s,* and

$ER_{i,t}$ = economic rent of unit *i* in period *t.*

The OER estimator uses the change in the pure rent which excludes the cost of any utilities included in the rent contract. In a parallel calculation to the rent estimator, the sum of the current pure rents for sampled, renter-occupied units within a segment, weighted by the owner weights, is divided by the sum of the weighted pure rents 6 months earlier.

This ratio is used to represent the 6-month change in the OER index for all owner-occupied housing units in the segment:

$$RER^{OER}_{t-6,t,s} = \frac{\sum_{i \in s} PR_{i,t} \cdot OW_s}{\sum_{i \in s} PR_{i,t-6} \cdot OW_s}$$

where

$RER^{OER}_{t-6,t,s}$ = relative OER price change between months *t* and *t–6* in segment *s,*

t = current period,

$OW_s =$ owners' equivalent rent weight for segment *s,* and

$PR_{i,t} =$ pure rent for unit *i* in period *t.*

The 6th root of the $REL_{t-6,t}$ is calculated to provide 1-month price relatives for index estimation:

$$REL_{t-1,t} = \sqrt[6]{REL_{t-6,t}}$$

Vacancy imputation

Vacant units that were previously occupied by renters are used in the calculation of relatives. The vacancy imputation process incorporates several assumptions about the unobserved rents of vacant units. It is presumed that rents tend to change at a different rate for units that become vacant (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 occupants or a unit becomes vacant, the rent is assumed to jump at some 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 nonvacant sample units in the PSU which have had a change in tenant during the 6-month period between *t–6* and *t*. Rent changes for nonvacant 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 in period *t* from their rent in period *t–6*.[1]

$r_{i,t} = r_{i,t-7}$ if the unit was not vacant in *t–6*, or

$r_{i,t} = r_{i,t-7} \cdot e^{J-C}$ if the unit was vacant in *t–6*,

where

$r_{i,t} =$ imputed rent of vacant rental unit *i* in period *t,*

J = the 6-month jump rate calculated for the PSU, and

C = the 1-month steady rate of change.

The imputation of vacant rents ensures that the unobserved rent change that occurs when a unit becomes vacant is reflected in the final index for rent. The 6-month rent-change estimates capture these changes once the units become occupied.

Non-interview imputations

Housing units that were previously responding but not currently responding and not vacant are also imputed and used in the calculation of the 1-month and 6-month relatives. All units within a PSU are broken up into high, medium, and low rent categories based on their rent level in *t–6*. The rents of nonresponding, nonvacant units are imputed forward into *t* by using the average rent change of other housing units in their respective category.

Aging adjustment

The aging adjustment accounts for the small loss in quality as housing units age (or depreciate) between interviews. The aging adjustment factors are $\frac{1}{(1-d)}$ where d is the monthly rate of physical depreciation. BLS computes factors for each housing unit using a multinomial logistic regression that controls for the age of the unit and a number of structural characteristics.[2]

Special pricing for seasonal items

Seasonal items are those commodities and services that are available only at certain times of the year rather than year round. Down parkas, baseball tickets, and bathing suits are examples of seasonal items. Special procedures are employed when selecting and pricing items generally available only part of the year to ensure that they are appropriately represented in the sample and that price changes are correctly included in the calculation of the CPI. In particular, the procedures prevent replacing a seasonal item when it is out of season.

Although seasonal items can exist in any ELI, some ELIs include an especially large percentage of such items and, consequently, receive special treatment. These seasonal ELIs include most apparel items and admission to sporting events. The designation of an ELI as seasonal or nonseasonal is made at the regional level, using the four geographic census regions in the CPI design. Some items that exhibit a seasonal selling pattern in the Northeast region, for example, may be sold year round in the South. In practice, though, nearly all ELIs designated seasonal are seasonal in all four regions.

After the samples for these seasonal ELIs are selected following the normal sample selection procedures, the number of quotes is doubled. This doubling ensures that, despite the seasonal disappearance of a substantial number of quotes, a large enough number of in-season quotes remains to calculate the index.

The quotes in these ELIs are paired. For each original quote that is selected, a second quote in the same ELI and outlet is initiated and priced 6 months later. 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 because these are the distinctions that are most commonly used by the retail trade 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.

Data collectors attempt to price every item in each period during which 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. A common practice in marketing seasonal items, particularly seasonal clothing, is to mark down prices to clear the merchandise from the stores as the end of each season approaches. During the period when a seasonal item is unavailable, its price is imputed following standard imputation procedures. When an item returns at the beginning of its season several months later, the price is directly compared with the item's last price, as it has been imputed forward. This completes the circle in a sense: having followed the price of the item down to clearance price levels, BLS then follows the price back up to regular (or at least higher) prices the following season. (Keep in mind that, in this context, the "following" season means the same season the next year;

that is, the following fall/winter season for the fall/winter sample, and the following spring/summer season for the spring/summer sample.)

When an item becomes permanently unavailable, the standard procedure is to replace it with the most similar item sold in the outlet. In the case of a year-round item that is not in a seasonal ELI, this process takes place as soon as the item is permanently unavailable. For items that are in seasonal ELIs and seasonal items in ELIs that are not designated seasonal, however, the period during which a replacement 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 that the correct balance of seasonal and year-round items is maintained. As a result, the estimates of price movement for the ELIs that include seasonal items correctly reflect price changes not just for items available year round but for the entire universe of items included in those ELIs.

Other price adjustments and procedures

There are many circumstances which can cause the price paid for goods and services by consumers to differ from the price posted. Where possible, BLS data collectors attempt to account for these circumstances, using a variety of procedures.

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 offer ounces of toothpaste free with the purchase of the regular 6-ounce tube, the item’s price is adjusted to reflect a decrease in the price per ounce. When the bonus is removed, the price per ounce returns to its previous level, and a price increase is recorded. In this instance, the value to the consumer is assumed to be one-third greater during the bonus period. If the bonus merchandise consists of an item that has some significant value to the consumer, and the item is different, an adjustment is made to account for the value of the free item when it is feasible to do so. Bonuses that are contingent on an additional unrelated purchase, such as a free can of soup when purchasing a whole chicken are ignored.

Cents-off coupons

For a coupon to be used to reduce the reported price of an item, the coupon must be either attached to the item, attached to the product’s display shelf, dispensed by machines attached to the product’s display shelf, located at promotional displays, or distributed to all shoppers by product representatives standing in the immediate vicinity of the display shelf. All other coupons presented by customers as purchase reductions at the time of payment are ineligible.

Concessions

A concession is a deduction of a specific amount from the proposed selling price for the item. The usual CPI practice is to subtract the average concession for the priced item over the past 30 days from the proposed selling price.

Container deposits

BLS collects information on container deposits for a variety of nonalcoholic and alcoholic beverages to reflect the influence of changes in deposit legislation on price change. Consumers who purchase throwaway containers are considered to be purchasing both the product itself and the convenience of throwing the container away. When a local jurisdiction enacts deposit 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 this convenience. The price of the same-sized container of product plus its deposit establishes an upper bound for the price change, because 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 legislation can be used to estimate price change when a container bill is repealed. Changes due to the enactment or repeal of container-deposit bills are shown in data for the month in which the legislation becomes effective.

Different-day pricing

For a subset of items, if the priced item that has been selected is not available for sale at the time of collection, prices from up to 7 days prior to the actual day of collection are eligible. The item must have been offered for sale during the previous 7 days and the most recently available price is reported. The list of eligible items generally consists of specific items that may not be available every day, such as a specific type of fresh fish.

Discounts

A discount price is a reduced price that is available only to certain customers in a specific outlet. If the discount is available only during the period of price collection, such as a back-to-school discount, the discount is included only if 50 percent or more of sales for the affected item are discounted. If the discount is in effect for more than one collection period and the discount applies to 5 percent or more of the dollar sales of the item in the outlet, a probability selection is made to determine which price should be collected. For example, if the regular cash price accounts for 84 percent of sales, senior citizens' discounts account for 10 percent and employee discounts account for 6 percent of sales, a one-time probability-based selection is made among the three options to determine which price to report.

Manufacturers' rebates

When product manufacturers offer customers cash rebates at the time of purchase for items priced in the CPI, these rebates are reflected in the index as price reductions. When a rebate is offered for a priced new vehicle, it is the estimated average rebate over the past 30 days that is subtracted from the vehicle's reported price. For vehicle leasing, it is the rebate in effect as of the day the collected price is obtained. For mail-in rebate offers, the price of the affected item is reported without subtracting the amount of the rebate. An attempt is made to determine the proportion of customers who take advantage of the rebate, and prior to its use in the index, the reported price is then adjusted accordingly.

Membership retail outlets

Outlets that require a membership fee to be paid in order to be able to shop at the outlet are eligible for pricing in the CPI. If the actual price paid for products varies with the level of membership, a specific membership is selected and the reported prices reflect that membership level.

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.

Sales taxes

The CPI includes all applicable taxes paid by consumers for services and products purchased. Some 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. Examples are tires and cigarettes. 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, county, and local tax structure governing the sale of the service or product at the point of purchase.

Shoppers' cards

If a priced outlet issues a card offering a card discount on selected products purchased by cardholders, such discounts are treated as temporary discounts and processed as follows. The discount is included only if signing up for the card is free and can be done by the consumer on the day of purchase.

Special-day prices

If a selected outlet has different prices for priced items based on the day of the week when a purchase is made, a selection is made between special-day and regular-day purchases, based on revenue. If the "special day" is selected, the price collected is for the most recent special-day price.

Unit-priced food items

When food items that are sold on a unit basis but lack a labeled weight are being priced, two items are weighed to permit calculation of an average weight for the item. This helps reduce the variability in size that occurs among individual, loose 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 report the price of one apple and the combined weight of two apples taken from the produce bin. In computing the price per ounce, the combined weight is divided by 2, and the 50-cent price of the Red Delicious apple is divided by this average weight.

Utility refunds

Sometimes, public utility commissions require that utilities such as telephone, natural (piped) gas, or electricity companies issue rebates to their customers for a number of different reasons. 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 CPI does not always view such refunds as reflecting current period prices for utility services. If all customers, both new and existing, are subject to having the refund applied to their bill, then the refund is included in the total price calculation. However, if the refund is only applied to those customers who were originally subject to the overcharge (i.e., existing customers only) then the refund is excluded. This procedure reduces the month-to-month volatility of utility indexes and ensures that they reflect current prices and price trends more accurately. Also excluded are refunds that are paid directly to consumers in a separate check and are not part of the bill. The utility indexes do include current-period credits that are based on current consumption, such as purchased gas adjustments and fuel adjustments.

Index calculation

As stated earlier, the CPI is actually calculated in two stages. Earlier sections described the first stage of that calculation: how the CPI calculates the basic indexes, which show the average price change of the items in each of the 7,776 CPI item-area combinations. The next section describes the second stage of calculation: how the aggregate indexes are produced by averaging across the 7,776 CPI item-area combinations.

Estimation of upper level price change

Aggregation of basic CPI data into published indexes requires three ingredients: basic indexes, basic expenditures to use as aggregation weights, and a price index aggregation formula that uses the expenditures to aggregate the sample of basic indexes into a published index.

Input basic price indexes

The CPI-U, CPI-W, and initial, interim, and final versions of the Chained CPI-U are constructed by using the same combination of modified Laspeyres and geometric mean basic indexes. In other words, the prices for each series are combined in the same way to form the basic price indexes.

CPI-U and CPI-W: input basic expenditure weights

In the CPI-U and CPI-W, aggregating basic indexes into published indexes using a modified Laspeyres formula requires an aggregation weight for each item-area combination. The function of the aggregation weight is to assign each basic index a relative importance or contribution in the resulting aggregate index. The aggregation weight corresponds to consumer tastes and preferences and resulting expenditure choices among the 243 basic items in the 32 basic areas comprising the CPI sample for a specified period.

Aggregation weights (AW) are defined as:

where

$_{i,a,p}P_{\varphi}$ = the estimated price of item *i* purchased in area *a* by population *p* in period φ, and

$_{i,a,p}Q_{\beta}$ = the estimated quantity of item *i* purchased in area *a* by population *p* in period β.

The period φ is the base period of the corresponding basic item-area index. For example, the “Sports equipment” (ITEM = RC02) for Seattle-Tacoma-Bellevue, WA (AREA = S49D) index series has a base period of φ = June 1985. CPI basic indexes have varying base periods, but most published indexes have an index base period of φ = 1982–84.

The period β corresponds to the reference period of the expenditures used to derive the implicit quantity weights needed for Laspeyres aggregation. As of 2018, the CPI-U and CPI-W had an expenditure reference period of β = 2015–16. BLS uses a biennial rotation schedule for updating the expenditure reference period. Effective with the January 2016 index, the expenditure reference period changed from β = 2011–12 to β = 2013–14; effective with the January 2018 index, it was updated again to 2015–16; and so forth. It is worth noting that a change in the expenditure reference period results in a change in the implicit quantity *Q* assigned to each basic index, but not the implicit price component *p* of the aggregation weight AW of each basic index.

Aggregation weights for the CPI-U and CPI-W are derived from estimates of household expenditures collected in the CE. Expenditure estimates at the basic item-area level would be unreliable due to sampling error without the use of statistical smoothing procedures. BLS uses two basic techniques to minimize the variance associated with each basic item-area base-period expenditure estimate. First, data are pooled over an extended period in order to build the expenditure estimates on an adequate sample size. The current reference period β uses 24 months of data.[3] Second, basic item-area expenditures are averaged, or composite estimated, with item-regional expenditures.[4] This has the effect of lowering the variance of each basic item-area expenditure at the cost of biasing it toward the expenditure patterns observed in the larger geographical area. This process is summarized in the equations in exhibit 2.

Exhibit 2. Estimation of CPI-U basic aggregation weights

$_{i,a,p,\beta_n}$ = expenditure on item *i* in area *a* by population *p* in year β_n

$_{\Sigma_{i,a,p,\beta_n}}$ = total expenditures in area *a* by population *p* in year β_n

$_{i,a,p,\beta_n}$ = share of total expenditures for item *i* in area *a* for population *p* in year β_n

= expenditure on item *i* in major area *m* by population *p* in year β_n

total expenditures in major area *m* by population *p* in year β_n

= share of total expenditures for item *i* in major area *m* for population *p* in year β_n

= composite-estimated share of total expenditures for item *i* in area *a*

for population *p* in year β_n

= estimated expenditure on item *i* in area *a* by population *p* in year β_n

= raked expenditures on item *i* in area *a* by population *p* in year β_n

= estimated expenditure in expenditure reference period β

= cost weight in pivot month *v*

= aggregation weight

where

P= population (all urban consumers or urban wage-earners and clerical workers);

a= CPI basic area;

i = CPI basic item;

e= expenditure class;

m= one of eight major areas, defined by census region and city-size classification (self-representing and non-self-representing);

p= price;

Q= quantity;

N= number of years in the CPI-U expenditure reference period (currently, N = 2);

β_n= year belonging to expenditure reference period β;

β= the reference period of the expenditures used to derive the implicit quantity weights for aggregation;

δ= weight assigned to major area *m,* where $0 \le \delta \le 1$;

φ= lower level index base period;

v = year and month, usually December, prior to the month when expenditure weights from the reference weights from reference period β are first used;

$_{i,a,p}S_{\beta_n}$= estimated expenditures *PQ* for item *i* in area *a* for population *p* in area *a* in period β_n;

$_{i,a,p}IX_{\varphi,\beta}$= lower level index of price change from index base period φ to expenditure reference period β for item *i* in area *a;* and

$_{i,a,p}IX_{\varphi,v}$ = lower level index of price change from index base period φ to pivot month *v* for item *i* in area *a.*

The estimated expenditure $_{i,a,p}(PQ)_\beta$ for item *I* in area *a* for population *p* in reference period β is derived from a weighted average of the item's relative importance in the basic area *a* and its relative importance in its corresponding region-size classification *m*, for each year encompassing reference period β. The weight δ

assigned to region-size class m and the weight $1-\delta$ assigned to the basic area a are a function of the variance in each area and the covariance of each measure.[5] The resulting average share $\hat{S}$ is then multiplied by the sum of all expenditures in the basic area in the corresponding year to obtain a revised item expenditure. In a process called raking, the revised item expenditures are adjusted by a factor such that, once summed, they equal the unadjusted expenditures at the region-size class m expenditure class e level. Annual item-area expenditures in year β_n have a lower bound of $0.01. The raked item expenditures in each year of reference period β are then averaged to obtain the aggregation weight: an expenditure value with an implicit price of period φ and implicit quantity of period β.

Initial C-CPI-U and interim C-CPI-U

The initial version of the C-CPI-U is published simultaneously with the CPI-U, so it uses expenditure data from the same expenditure reference period β as the CPI-U for its aggregation weights. Since 2015, BLS has issued four preliminary estimates of the C-CPI-U, by quarter, with final data being published approximately 1 year after the reference month. Hence, if the ensuing year was one in which the weight was updated, then the interim version of each monthly C-CPI-U was based on more contemporaneous expenditures than its initial version. For example, 2015 initial indexes produced in 2015 used β = 2011–12. However, 2015 interim indexes produced in 2016 were constructed using β = 2013–14.

Final C-CPI-U

For the C-CPI-U, which uses the Törnqvist index formula for upper level aggregation in a monthly chained construct, monthly expenditure estimates for each basic item-area combination are required as aggregation weights. These are derived from the same CE data as the CPI-U aggregation weights. Like the biennial data used for CPI-U aggregation, adequacy of the underlying sample size from which the expenditure weights are estimated is an issue for C-CPI-U aggregation. To minimize the variance of the basic item-area monthly expenditures, a ratio-allocation procedure is adopted to estimate each item-area monthly expenditure from U.S. monthly item expenditures.

Estimation of monthly expenditures at the basic level

Estimated monthly expenditures are given by

where

p = population (note that C-CPI-U is produced only for the all urban consumers population),

a = CPI basic area,

i = CPI basic item,

A = all CPI basic areas (U.S. city average),

P = price,

Q = quantity,

t = month, and

T = period covering month *t* and 11 months prior to month *t*.

The monthly expenditure for an item in a basic area is derived in two steps. First, the monthly expenditure for the item is summed across all 32 areas to obtain a U.S. monthly item expenditure. Second, the U.S. monthly item expenditure is allocated among all 32 basic areas, according to each area's relative expenditure share for the item during the current and preceding 11 months. Note that:

The estimated monthly item-area expenditures have a lower bound of $0.000833 (1/12 of a cent), and when summed over the calendar year, they have a lower bound of $0.01, which is equivalent to that of the annual data in the CPI-U expenditure reference period.

Aggregation formula

A modified Laspeyres price index is used to aggregate basic indexes into published CPI-U and CPI-W indexes.

The Laspeyres index uses estimated quantities from the predetermined expenditure reference period β to weight each basic item-area index. These quantity weights remain fixed for a 2-year period, and then are replaced in January of each even-numbered year when the aggregation weights are updated. In a Laspeyres aggregation, consumer substitution between items is assumed to be zero. The aggregate index for any given month is computed as a quantity-weighted average of the current month's index divided by the index value in the index base period. Month-to-month price change is then calculated as a ratio of the long-term monthly indexes. The relevant equations follow.

CPI-U and CPI-W upper level aggregation formula

Long-term price change is given by

Month-to-month price change is given by

where

A = all basic areas (U.S. city average),

a = CPI basic area,

p = populations (all urban consumers or urban wage earners and clerical workers),

i = CPI basic item,

I = all basic items,

t = month,

z = base period of the aggregate index (the CPI-U U.S. city average index series for all items has a base period of 1982–84),

φ = base period of the basic index for item i in area a,

β = the reference period of the expenditures used to derive the implicit quantity weights needed for aggregation,

v = pivot month (usually December) prior to the month when expenditure weights from period β are first used in the CPI,

$_{ap}I_{ai}^{\varphi t}$ = lower level index of price change from period φ to month t for item i in area a for population p,

$_{ap}I_{ai}^{\varphi v}$ = lower level index of price change from period φ to pivot month v for item i in area a for population p,

$_{ap}AW_{\beta}$ = aggregation weight from reference period β for item i in area a for population p,

$_{ap}I_{zv}$ = aggregate level CPI series of price change from period z to pivot month v for aggregate area i in aggregate area a for population p,

L = index calculated using a Laspeyres formula, and

G = index calculated using a geometric mean formula.

In contrast, the C-CPI-U is built by chaining together indexes of 1-month price changes. For the final C-CPI-U index, each monthly index is computed using the Törnqvist formula with monthly weights from both the current and the previous month. Consumer substitution behavior is not assumed by the Törnqvist formula; rather, it is implicitly accounted for by use of current- and base-month expenditures. An index of 1-month price change is calculated

and then multiplied by the index value for the previous month to obtain the current-month index value. Following are the relevant equations.

Final C-CPI-U upper level aggregation formula

Long-term price change is given by

$$[illegible]$$

and month-to-month price change is given by

$$[illegible]$$

where

A = all basic areas (U.S. city average),

a = CPI basic area,

i = CPI basic item,

I = all basic items,

t = month,

z = base period of the aggregate index (the C-CPI-U U.S. city average index series for all items has a base period of December 1999),

φ = base period of the basic index for item i in area a,

$[illegible]$ = lower level index of price change from period φ to month t for item i in area a,

$[illegible]$ = lower level index of price change from period φ to month $t–1$ for item i in area a,

$[illegible]$ = expenditure in month t for item i in area a as a percentage of total expenditures in month t for aggregate item i in aggregate area A,

$[illegible]$ = expenditure in month $t–1$ for item i in area a as a percentage of total expenditures in month $t–1$ for aggregate item i in aggregate area A, and

$[illegible]$ = aggregate level C-CPI-U Törnqvist index of price change from period z to month t for aggregate item i in aggregate area A.

BLS revises the C-CPI-U quarterly, using the constant elasticity of substitution formula for the calculation of the preliminary versions of that index. The initial version of the C-CPI-U is released concurrently with the CPI-U for each calendar month. The final version of the index is released approximately 10-12 months later. In between the initial release and the final release, there are three quarterly updates. The 1-month price change for each interim release is the same as the initial version. The interim versions reflect only updates to index levels—that is, the value of the index in a given month relative to the value in its base period. These updates result from the conversion of 1-month price changes from initial to final value in preceding months in the monthly chained series. The constant elasticity of substitution uses an estimate of consumer substitution that lies between the estimates assumed in the geometric mean and Laspeyres formulas, and represents a model that is closer to actual consumer behavior. This estimate of consumer substitution σ is called the elasticity of substitution. For additional information on the C-CPI-U framework, see the article [Improving initial estimates of the Chained Consumer Price Index](#).

Month-to-month price change under the constant elasticity of substitution formula is given by:

$$[illegible]$$

The constant elasticity of substitution pivoted expenditure weight for a biennial period is given by:

$$[illegible]$$

where

A = all basic areas (U.S. city average),

a = CPI basic area,

i = CPI basic item,

I = all basic items,

t = month,

b = biennial expenditure reference period,

x = index base period (initially December 1999 = 100),

V = pivot month,

$P_b^{i,a}$ = price of item i in area a during period b,

$Q_b^{i,a}$ = quantity of item i in area a during period b,

σ = elasticity of substitution for the index period, and

$IX_{i,a,t}$ = lower level index for item i in area a in month t.

Calculation of seasonally adjusted indexes

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, and sales. 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 X-13ARIMA-SEATS seasonal adjustment software. In some cases, intervention analysis seasonal adjustment is carried out using X-13ARIMA-SEATS to derive more accurate seasonal factors. Consumer price indexes may be adjusted directly or aggregately, depending on the level of aggregation of the index and the behavior of the component series.[6]

Intervention analysis and seasonal adjustment

Some index series show erratic behavior due to nonseasonal economic events (called interventions) or methodology changes. These events, which can be one-time occurrences or recurring events that happen at infrequent and irregular intervals, adversely affect the estimate of the seasonal component of the series.

Intervention analysis seasonal adjustment allows nonseasonal economic phenomena, such as outliers and level shifts, to be factored out of indexes before calculation of seasonal adjustment factors. (An outlier is an extreme value for a particular month. A level shift is a change or shift in the price level of a CPI series caused by an event, such as an excise tax increase or oil embargo, occurring over 1 or more months.) An index series whose underlying trend has experienced a sharp and permanent shift will generate distorted results when adjusted using the standard X-13ARIMA-SEATS procedure. The X-13ARIMA-SEATS regression techniques are used to model the distortions and account for them as part of the seasonal adjustment process. The result is an adjustment based on a representation of the series with the seasonal pattern emphasized. Intervention analysis seasonal adjustment also makes it possible to account for seasonal shifts, resulting in better seasonal adjustment in the periods before and after the shift occurred. Not all CPI series are adjusted using intervention analysis seasonal adjustment techniques. These seasonal factors are applied to the original unadjusted series. Level shifts and outliers, removed in calculating the seasonal factors, remain in the resulting seasonally adjusted series.

In recent years, BLS has used intervention analysis seasonal adjustment for various indexes, such as gasoline, fuel oil, new vehicles, women’s and girls’ apparel, educational books and supplies, electricity, utility (piped) gas service, water and sewerage maintenance, nonalcoholic beverages and beverage materials, and whiskey at home. Series are adjusted using intervention analysis techniques when interventions are clearly identified. After a number of years, series may revert to adjustment using standard methods. For some series, intervention analysis is used and the resulting series does not show a clear and stable seasonal pattern. In these cases, the series is not seasonally adjusted.

Direct and aggregative adjustment

Each year, BLS seasonally adjusts eligible lower-level CPI index series directly with the X-13ARIMA-SEATS software using unadjusted indexes for the latest 5 to 8 calendar years. CPI index series are adjusted using the multiplicative model. Most high-level index series are adjusted by the 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 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 that have not been selected for seasonal adjustment, the original unadjusted data are used in the aggregation process.

Revision

The seasonal factors are updated annually. Each year in February, BLS recalculates and publishes revised seasonally adjusted indexes for the previous 5 years. Seasonally adjusted indexes become final in the 5th and last year of revision. Seasonal factors for the past year are used to generate seasonally adjusted indexes for the current year starting with the release of the January CPI.

Calculation of annual and semiannual average indexes

CPI annual average indexes use 12 successive months of CPI values:

$$I_{avg} = \frac{\sum_{t=1}^{12} I_t}{12}$$

Semiannual average indexes are computed for the first half of the year (January to June) and for the second half of the year (July to December) using 6 successive months of CPI values:

$$I_{avg} = \frac{\sum_{t=1}^{6} I_t}{6}$$

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

Average prices

Average prices are estimated from CPI data for selected food and beverage items, utility (piped) gas, electricity, gasoline, automotive diesel fuel, and fuel oil number 2 to support the research and analytic needs of CPI data users. (See appendix 6.) Average food prices are published without tax, while the other average prices are published with tax included.

All eligible prices are converted to a price per normalized quantity. These prices are then used to estimate a price for a defined fixed quantity. For example, prices for a variety of package sizes for flour are converted to prices per ounce. An average price per ounce of flour is then estimated and multiplied by 16 to yield a price per pound, the published quantity.

The average price for collection period t is estimated as

$$P_t = \frac{\sum_i W_{it} P_{it}}{\sum_i W_{it} P_{ib}}$$

where

W_{it} = the quote-level expenditure weight of items used in the average price estimation for the ELI/PSU/replicate in time period *t*,

P_{ib} = the price of item *i* in the base period,

P_{it} = the price of item *i* in period *t,* and

P_t = the average price for period *t.*

Dividing the expenditure weight by the base price 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. Imputed prices are used in estimating average prices.

Precision of CPI 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 (replications) 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.

Sources of error

We divide the total error into two sources: sampling error and nonsampling error. Sampling error is the uncertainty in the CPI caused by the fact that a sample of retail prices is used to compute the CPI, instead of using the complete universe of retail prices. The sampling variance attributable to the estimation of expenditure weights is not directly incorporated in the variance estimates computed for the CPI.[7] Research suggests that the impact of CE sample sizes is on the variance of the variance and not on the expected value of the variance of CPI estimates. Nonsampling error is the rest of the error, and will be discussed at the end of this section. Incorrect information given by survey respondents and data processing errors are examples of nonsampling error.

BLS constantly tries to improve the precision of the CPI. Variance and sampling error are reduced by using samples of retail prices that are as large as possible, given resource constraints. BLS has developed a model that optimizes the allocation of resources. The model indicates the number of prices that should be observed in each geographic area and each item category to minimize the variance of the U.S. city average all-items index. BLS reduces nonsampling error through a series of computerized and professional data reviews, as well as through continuous survey process improvements and theoretical research.

Sampling error

Starting in 1978, the CPI's sample design has accommodated variance estimation by using two or more independent samples of items and outlets in each geographic area. This allows two or more statistically independent estimates of the index to be made. The independent samples are called replicates, and the set of all observed prices is called the full sample.

As discussed earlier, BLS calculates indexes for 32 geographic areas across the United States. The 32 areas consist of 23 self-representing areas and 9 non-self-representing areas. Self-representing areas are large metropolitan areas, such as the Boston and the San Francisco metropolitan areas. Non-self-representing areas are collections of smaller metropolitan areas. For example, one non-self-representing area is a collection of 64 small metropolitan areas in the Middle Atlantic division (Pittsburgh, Buffalo, Rochester, Reading, and others) of which four metropolitan areas have been randomly selected to represent the entire set. Within each of the 32 areas, price data are collected for 243 basic item categories. Together, the 243 basic item categories cover all consumer purchases.

Multiplying the number of areas (32) by the number of item strata (243) gives 7,776 different item-area combinations for which price indexes need to be calculated. Separate price indexes are calculated for each one of these 7,776 item-area combinations. After calculating all 7,776 of these basic level indexes, the indexes are then aggregated to form higher level indexes, using expenditure estimates from the CE as their weights.

CPI variances are primarily computed with a stratified random groups method, for 1-, 2-, 6- and 12-month percent changes. Since 1998, BLS uses the stratified random groups method, in which replicate percent change estimates are computed separately for certain subsets of areas by substituting replicate cost weights for full sample cost weights, and then those individual percent change estimates are subtracted from the full sample percent change estimate and squared. These estimates are combined to produce the variance of the entire item-area combination.

Variance estimation using replicates

Let IX*(A,I,f,t)* denote the index value for area *A*, item category *I*, in month *t*, where *f* indicates that it is the full sample value, and let IX*(A,I,f,t–k)* denote the value of the same index in month *t–k*. The uppercase letter *A* denotes a set of areas, such as the Northeast or Midwest region of the country, and the uppercase letter *I* denotes a set of item strata, such as all items or all items less food and energy, or a single item stratum. Also, let IX*(A,I,r,t)* and IX*(A,I,r,t–k)* be the corresponding index values for replicate *r*. Most areas have two replicates, but some have more.

Then the full-sample *k*-month percent change between months *t–k* and *t* is computed by dividing IX*(A,I,f,t)* by IX*(A,I,f,t–k)*, subtracting 1, and multiplying by 100:

$$P(A,I,f,t-k) = \left(\frac{IX(A,I,f,t)}{IX(A,I,f,t-k)} - 1\right) \times 100$$

where

$P(A,I,f,t-k)$ = full-sample *k*-month percent change between months *t–k* and *t,* for item category *I* in area *A;*

IX*(A,I,f,t)* = index value for area *A,* item category *I*, for the full sample, in month *t;*

IX*(A,I,f,t–k)* = index value for area *A,* item category *I,* for the full sample, in month *t–k;*

A= area;

I = item category;

t = month;

t–k = month *k* months before month *t;*

f = subscript indicating that the value is for the full sample; and

r = parameter indicating that the value is for a replicate.

Every index has an aggregation weight AGGWT*(A, I, f)* or AGGWT*(A, I, r)* associated with it, which is used to combine the index with others to produce aggregate indexes for larger geographic areas and larger item categories. For example, the aggregation weights are used to combine all 7,776 basic-level indexes into higher level indexes such as the U.S. city average all-items index.

The product of an index and its weight is called a cost weight:

where

CW*(A,I,f,t)* = cost weight for area *A,* item category *I,* for the full sample, in month *t;*

IX*(A,I,f,t)* = index value for area *A,* item category *I,* for the full sample, in month *t;*

AGGWT*(A,I,f,t)* = aggregation weight for area *A,* item category *I,* for the full sample, in month *t;*

A= area;

I = item category;

t = month; and

f = parameter indicating that the value is for the full sample.

A cost weight is an estimate of the total cost in area *A* for consumption of item category *I* in month *t*. Replicate cost weights are produced from replicate level indexes and full sample aggregation weights. Because the

aggregation weights are not indexed by time (except across pivot months; see the section below, “Bridging across pivot months”), the preceding percent change formula is equivalent to:

$$P(A,I,t,t-k) = \frac{CW(A,I,t)}{CW(A,I,t-k)} - 1$$

where

$P(A,I,t,t-k)$ = full-sample *k*-month percent change between months *t–k* and *t,* for item category *I* in area *A;*

$CW(A,I,t)$ = cost weight for area *A,* item category *I,* for the full sample, in month *t;*

$CW(A,I,t-k)$ = cost weight for area *A,* item category *I,* for the full sample, in month *t–k;*

A = area;

I = item category;

t = month;

t–k = month *k* months before month *t;* and

f = parameter indicating that the value is for the full sample.

which is equivalent to:

$$P(A,I,t,t-k) = \frac{CW(A,I,t)}{CW(A,I,t-k)} - 1$$

because cost weights are additive from the lowest area-item level up to the highest U.S. city average all items level. The lowercase letter *a* denotes 1 of the 32 basic-level areas included in area = *A*, and the lowercase letter *i* denotes 1 of the 243 item categories. (Note: Item aggregation *I* can be as small as one item stratum or may comprise one or more major groups.)

For the Stratified Random Groups method used here, replicate percent changes are defined as follows: full sample cost weights are used for every geographic area within area = *A* except for one of the areas. In the omitted area, the full sample cost weight is replaced by a replicate cost weight. Let the lowercase letter *a* denote 1 of the 32 basic-level areas included in area = *A*.

Then, the replicate percent change, for area = *a*, item = *I*, replicate = *r*, between months *t–k* and *t*, is computed as:

$$P(a,I,r,t,t-k) = \frac{CW(a,I,r,t)}{CW(a,I,r,t-k)} - 1$$

where

$$CW(a,I,r,t) = \sum_{a \neq a} CW(a,I,t)$$

$P(t,k,l,a,r)$ = replicate *k*-month percent change between months *t–k* and *t,* for item category *l* in area *A;*

a = basic area;

A = aggregate area;

l = item;

r = parameter indicating that the value is for a replicate;

f = parameter indicating that the value is for the full sample;

t = month;

t–k = month k months before month *t;* and

$CW(l,A,a)$ = cost weight in area *A* of item category *l*.

The variance is computed with the following stratified random groups variance estimation formula:

$$V(P(t,k,l,A)) = \frac{1}{2}\sum_{a=1}^{2}\left[P(t,k,l,a,f) - P(t,k,l,a,r)\right]^2$$

where

$V(P(t,k,l,A))$ = the variance of the full sample *k*-month percent change between months *t–k* and *t*, for item category *l* in area *A;*

$P(t,k,l,a,r)$ = replicate *k*-month percent change between months *t–k* and *t,* for item category *l* in area *A;*

$P(t,k,l,a,f)$ = full-sample *k*-month percent change between months *t–k* and *t,* for item category *l* in area *A;*

a = basic area;

A = aggregate area;

l = item;

r = parameter indicating that the value is for a replicate;

f = parameter indicating that the value is for the full sample;

t = month;

t–k = month k months before month *t;* and

R_a = the number of replicates in area *a.*

Finally, the standard error of the percent change is computed by taking the square root of its variance:

$$SE(PC_{A,I,(t-k)}) = \sqrt{V(PC_{A,I,(t-k)})}$$

where

$SE(PC_{A,I,(t-k)})$ the standard error of the full sample *k*-month percent change between months *t–k* and *t*, for item category *I* in area *A;*

$V(PC_{A,I,(t-k)})$ = the variance of the full sample *k*-month percent change between months *t–k* and *t*, for item category *I* in area *A;*

$PC_{A,I,(t-k)}^{r}$ = replicate *k*-month percent change between months *t–k* and *t,* for item category *I* in area *A;*

$PC_{A,I,(t-k)}^{f}$ = full-sample *k*-month percent change between months *t–k* and *t,* for item category *I* in area *A;*

a = basic area;

A = aggregate area;

I = item;

r = parameter indicating that the value is for a replicate;

f = parameter indicating that the value is for the full sample;

t = month; and

t–k = month *k* months before month *t.*

Variance estimation without replicates

BLS publishes index series for 82 special relative (item) categories (SRCs), which are below the item stratum level and thus do not have accompanying replicate index values. (CE weights are produced only down to the item-stratum level in each index area.) The CPI stratified random groups methodology requires a replicate structure. So, for these SRC items (such as butter or pork or new cars), an alternative variance estimation method is needed. Given the availability (at the regional and higher area levels) of independent estimates for these SRC items, the jackknife variance estimation methodology can be employed. Each area's full-sample cost weight can be subtracted from the all-area full-sample cost weight to provide a jackknife replicate estimate. By taking the ratio of these replicate cost weight estimates at times *t* and *t–k*, subtracting 1, and multiplying by 100, one obtains the

required jackknife replicate percent change value. (For the U.S. city average special item estimates, there are 32 independent index areas, and so there are 32 jackknife replicate estimates with which to work.)

The full-sample percent change is computed as before (except that, here, item category = *I* is smaller even than an item stratum):

$$P(A,I,t,t-k) = \frac{CW(A,I,t)}{CW(A,I,t-k)} - 1$$

where

$P(A,I,t,t-k)$ = full-sample *k*-month percent change between months *t–k* and *t,* for item category *I* in area *A;*

$CW(A,I,t)$ = cost weight for area *A,* item category *I,* for the full sample, in month *t;*

$CW(A,I,t-k)$ = cost weight for area *A,* item category *I,* for the full sample, in month *t–k;*

A = area;

I = item category;

t = month;

t–k = month *k* months before month *t*; and

f = parameter indicating that the value is for the full sample.

The jackknife replicate percent change is computed as follows:

$$P_{r}(A,I,t,t-k) = \frac{CW_{r}(A,I,t)}{CW_{r}(A,I,t-k)} - 1$$

Then the variance for the *k*-month percent change is computed in the usual jackknife form:

$$Var(P(A,I,t,t-k)) = \frac{32}{31} \sum_{r=1}^{32} \left[P_{r}(A,I,t,t-k) - P(A,I,t,t-k) \right]^{2}$$

Building across pivot months

Every 2 years, BLS updates its set of aggregation index weights based on CE data collected from the *t–2* and *t–3* years. In January 2016, BLS replaced its old set of aggregation weights with a new 2-year set of weights from expenditure data collected in 2013–14. In January 2018, this set of weights was replaced by an updated set of weights from expenditure data collected in 2015–16, and so on.

Whenever the variance estimates cross the pivot month (as they did in December 2015 and December 2017), a bridging factor has to be introduced into any variance calculation that crosses the pivot month anywhere between *t* and *t–k* months (including month *t–k*, but not including month *t*). The bridging factor is then applied directly to

the individual ratio of cost weights, for both full-sample and replicate values, inside each percent change calculation.

Thus, in its most general form:

$$\frac{CW(t,k,old)}{CW(t,k,new)}$$

for every combination of area and item, and for full-sample and replicate values, with the bridging factor defaulting to 1 whenever not applicable.

The bridging factor essentially allows the old aggregation weight in the bridge's numerator to cancel out the old aggregation weight in the *t–k* cost weight, while the new aggregation weight in the bridge's denominator cancels out the new aggregation weight in the *t* cost weight, leaving the index relative free to move this level's percent change without disruption.

where

$\frac{CW(t,k,old)}{CW(t,k,new)}$ = bridging factor,

$CW(t,k,old)$ = old cost weight,

$CW(t,k,new)$ = new cost weight,

t = month, and

$t-k$ = the month that is *k* months before month *t.*

Nonsampling error

Surveys involve many operations, all of which 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). Such errors arise from the omission of cities, households, outlets, and items that are part of the target populations from the relevant sampling frames or from their double-counting or improper inclusion in the frames. A potential source of coverage error is the time lag between the Consumer Expenditure Survey (CE) and the initiation of price collection for commodities and services at sampled 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 CE 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. Response rates during monthly pricing for the CPI C&S and housing surveys are published annually and available online.

Response error results from the collection and use of incorrect, inconsistent, or incomplete data during estimation. 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 employs an independent pricing methodology specifically to avoid potential response bias.

Processing error arises from incorrect editing, coding, and data transfer. Price data are collected by computer-assisted data collection. Automated data checking ensures that only correct data types are collected; other automated logic checks remove all redundant question patterns, and the instrument informs staff when not all required data have been collected. Errors can also result from software problems in the computer processing that cause correctly entered data to be lost. Computer screening and professional review of the data provide checks on processing accuracy. 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 it is intended to measure. 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.

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. Because 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 underestimates quality change for particular kinds of items.

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. In 1988, BLS began adjusting for the slow depreciation of houses and apartments over time. BLS research indicates that annual changes in 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 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 the number of eligible sample units and 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 the number of eligible sample units and the number of sample units with eligibility not determined.

Commodities and services items (except rent and owner's equivalent rent) are further broken down into outlets and quotes. An outlet is a generic term used 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. Relatively low percentages of quotes are reported collected and reported used in apparel estimation. Low rates for these items can mostly be attributed to the design of the apparel sample. Because apparel items are commonly in stores only at 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 many apparel quotes, although eligible, are designated "out of season," and prices are not collected. For additional information, see the earlier subsection on seasonal items.

The response rates for housing (shelter) include categories for rental units only; owner-occupied housing units are out of scope for the CPI housing sample. A unit qualifies for inclusion if its tenure status is known either by previous knowledge or is collected in the current interview period. 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 eligible and the data are reported. If the assigned unit is located but is unoccupied, the unit is designated "eligible, found vacant." In instances where the unit is eligible but no data are available (for example refusals), the unit is designated "eligible, other." The response rates at the estimation phase are units that are used in either rent or rental equivalence.

NOTES

[1] For more information on vacancy imputation, see J.P. Sommers and J.D. Rivers, "Vacancy imputation methodology for rents in the CPI," *Proceedings of the American Statistical Association, Business and Economic Statistics Section* (Alexandria, VA: American Statistical Association, 1983).

[2] For further information, see Walter F. Lane, William C. Randolph, and Stephen A. Berenson, "Adjusting the CPI shelter index to compensate for effect of depreciation," *Monthly Labor Review*, October 1988, pp. 34–37.

[3] Prior to 2002, the expenditure reference period was based on 36 months of data (for example, β = 1993–95 from 1998 to 2001 and β = 1982–84 from 1987 to 1997).

[4] Basic areas are grouped into city-size classifications by region for the purpose of composite estimation. There are four regions (Northeast, Midwest, South, and West) and two city-size classifications (A-sized cities and non-A-sized cities) for a total of eight regional city-size classifications.

[5] For more information on composite estimation, see Michael P. Cohen and John P. Sommers, “Evaluation of the methods of composite estimation of cost weights for the CPI,” Proceedings of the American Statistical Association, Business and Economic Statistics Section (Alexandria, VA: American Statistical Association, 1984.), pp. 466–471.

[6] J.A. Buszuwski and S. Scott, "On the use of intervention analysis in seasonal adjustment," Proceedings of the American Statistical Association, Business and Economics Section (Alexandria, VA, American Statistical Association, 1988).

[7] See the U.S. Bureau of Labor Statistics Consumer Expenditure Survey Methodology for more detail on consumer expenditure weights, https://www.bls.gov/opub/hom/cex/calculation.htm.

Last Modified Date: November 24, 2020

Presentation

The Consumer Price Index (CPI) is a measure of the average change over time in the prices paid by consumers for a representative basket of consumer goods and services. The CPI measures inflation as experienced by consumers in their day-to-day living expenses. The CPI is used to adjust income eligibility levels for government assistance, federal tax brackets, federally mandated cost of living increases, private sector wage and salary increases, and consumer and commercial rent escalations. Consequently, the CPI directly affects hundreds of millions of Americans.

The Consumer Price Index (CPI) is published monthly in a news release, along with supplemental tables and a database. Additional information is available in various analytical papers and factsheets.

The Bureau of Labor Statistics publishes four consumer price series each month:

- CPI for All Urban Consumers (CPI-U)
- CPI for Urban Wage Earners and Clerical Workers (CPI-W)
- Chained CPI for All Urban Consumers (C-CPI-U)
- Average prices

CPI-U

The CPI-U, which began publication in January 1978, represents the buying habits of the residents of urban or metropolitan areas in the United States and covers about 93 percent of the U.S. population. The CPI-U is the headline CPI number and is the one most commonly referred to in the news. Not seasonally adjusted data for this series is available back to 1913. Seasonally adjusted data for this series is available back to 1947.

CPI-W

The CPI-W, the oldest of the series, is computed using the same prices as the CPI-U, but the weights of the CPI-W are based on a subset of the CPI-U population. This subset covers about 29 percent of the population. The CPI-W population includes households where more than one-half of the household's income must come from clerical or wage occupations, and at least one of the household's earners must have been employed for at least 37 weeks during the previous 12 months. This information is obtained from the Consumer Expenditure Survey Interview Questionnaire. Thus, the CPI-W population excludes households of professional and salaried workers, part-time workers, the self-employed, and the unemployed, along with households with no one in the labor force, such as those of retirees. Not seasonally adjusted data for this series is available back to 1913. Seasonally adjusted data for this series is available back to 1947.

C-CPI-U (Chained CPI)

The C-CPI-U also represents the urban population as a whole, but uses a different formula and different weights to combine basic (or lower level) indexes. The formula used in the C-CPI-U accounts for changes in consumer spending behavior by using weights based on expenditures from both a base period and a current period. This formula requires consumer spending data that are not immediately available; therefore, the C-CPI-U, unlike the

other two series, is published first in preliminary form and is subject to three quarterly revisions. BLS began publishing this series in August 2002, with data starting in January 2000.

Average prices

For some food, beverage, and energy items, the CPI publishes average price data. A list of what is covered in the published average price series is shown in appendix 6. Average price data is available from 1980.

Item and area indexes

Item (a good or service in the market basket) and area (section of the United States) indexes use only a portion of the CPI sample; this makes them subject to substantially greater sampling error than the national CPI. The primary reason for publishing indexes at the item level is to aid in the analysis of movements in the national all items index. Decisions on which detailed indexes to publish depend, in part, on the reliability of their estimates.

Item indexes

BLS classifies the CPI market basket of consumer goods and services into a hierarchy of categories that are also published monthly. The top levels of the item category hierarchy consist of the following:

- Eight major groups: food and beverages, housing, apparel, transportation, medical care, recreation, education and communication, and other goods and services
- Other groups (sets of related indexes that are subsets of a major group but include multiple expenditure classes): e.g., private transportation
- Expenditure classes: e.g., motor fuel
- Item strata: e.g., gasoline, unleaded regular

BLS also publishes indexes for special aggregations, such as energy items, that cut across the preceding classification scheme. Users consider the series all items less food and energy to measure the core rate of inflation. Food and energy are two of the most volatile components of the CPI, so many analysts regard the measure of core inflation as more useful for their purposes. We publish all levels down to item strata at the U.S. level and less item detail for the area indexes.

Area indexes

Along with the U.S.-level data, we calculate and publish indexes for the following areas on a monthly or bi-monthly basis:

- 4 Census regions: Northeast, South, Midwest, West;
- 9 Census divisions: New England, Mid Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, Pacific;
- 2 population-size classes: A (greater than 2,500,000) and B/C (2,500,000 or less), for the U.S. and the 4 Census regions; and
- 23 selected metropolitan areas: Chicago, Los Angeles, and New York are published monthly; the remaining areas are published bi-monthly.

Detailed food, energy, and shelter indexes are published monthly. For a detailed list of items that are published monthly according to these areas, see appendix 7.

Interpreting an Index

Each month's index value displays the average change in the prices of consumer goods and services since a base period, which is currently 1982–84 for most indexes. For example, the CPI-U for April 2017 was 244.524. You can interpret this as a representative set of consumer goods and services that cost $100 in 1982–84 would have cost $244.52 in April 2017.

Rather than emphasizing the level of the index in comparison to the base period, the monthly CPI release stresses the CPI's percent change from the previous month and from the previous year. The most commonly reported monthly percent changes are the 1-month seasonally adjusted percent change, and the 12-month not seasonally adjusted percent change.

Correction and revision policies

CPI-U and CPI-W

The CPI does not rely on respondents to send data to the BLS national office; CPI data collectors gather almost all of the data for the CPI-U and CPI-W. Virtually all data are received in time for the calculation of indexes for the appropriate month, so routine revisions to account for late-arriving data are not necessary. The CPI-U and CPI-W are commonly used in escalation agreements and to adjust pensions and tax brackets; consequently, revisions can be costly for the users of these indexes so these series are final when issued. Additional information on escalation contracts is available in the How to Use the Consumer Price Index for Escalation factsheet. In the case when BLS discovers that an error was made while collecting or compiling information, we issue a correction to the affected series in accordance with BLS policies.

C-CPI-U

As previously noted, C-CPI-U indexes are not final when first issued; they are subject to three quarterly revisions. The data are final after the last revision, which occurs 10-12 months after the initial publication. If the CPI-U and CPI-W series are corrected, the C-CPI-U series is corrected as well for all series affected by the error, as far back as the previous 5 years.

Seasonal adjustment

Seasonal adjustment is a statistical procedure that adjusts for normal seasonal variations to help decisionmakers better understand underlying economic changes (e.g., gas prices that typically rise each summer). Each year, with the release of January data, seasonal adjustment factors are recalculated to reflect price movements from the previous calendar year. This routine recalculation will result in revisions to seasonally adjusted indexes for the previous 5 years. Additional information on seasonal adjustment is available on the seasonal adjustment webpage.

Uses of the CPI

The CPI affects virtually all Americans because of the many ways in which it is used. Its major uses include the following.

As an economic indicator

As the most widely used measure of retail inflation, the CPI is a major indicator of the effectiveness of government economic policy. The President, Congress, and the Federal Reserve Board use the movement of the CPI to help

formulate and monitor the effects of fiscal and monetary policy. Business executives, labor leaders, and other private citizens also use the index as a guide in making economic decisions.

As a means of adjusting income payments

The index directly affects the income of almost 80 million people. Social Security benefits as well as military and Federal Civil Service pension payments are indexed to the CPI. In the private sector, many collective bargaining agreements tie automatic wage increases to the CPI and some private firms and individuals use the index to keep rents, alimony, and child support payments in line with changing prices.

To adjust tax bracket thresholds

Federal (and some state) income tax bracket thresholds and other parameters are adjusted to the CPI. Additional information about this process is available on the Internal Revenue Service website.

As a deflator of other economic series

Other statistical programs use the CPI or its components to adjust for price changes to produce inflation-free versions of their series. Examples of CPI-adjusted series include components of the U.S. Department of Commerce's National Income and Product Accounts (such as gross domestic product and personal consumption expenditures), retail sales measures, and the BLS hourly and weekly earnings series.

Last Modified Date: November 24, 2020

History

Although the official Consumer Price Index (CPI) began in 1919, origins of the CPI date back about 30 years earlier. The first major study conducted by the new Bureau of Labor (later to be renamed the Bureau of Labor Statistics) that influenced later work on a “cost of living” index was an examination of family expenditures and retail prices between 1888 and 1890.[1]

In 1901, BLS began to systematically collect retail prices for food as part of a consumer expenditures and income study. Between 1908 and 1920, much discussion arose concerning the statistical methods used to measure price change and the commodity content of an index designed to measure changes in the cost of living or more specifically, the general consumer price level. In the period of rapidly increasing prices during and immediately following World War I, it became increasingly clear that a measure of the change in food prices was not an adequate measure of the cost of living nor the general price level. Arbitration boards and commissions were considering many aspects of living costs in rendering rulings and settlement awards, and their demands for data helped to shape the scope, concept, and procedures for the index.[2]

BLS conducted studies of family expenditures in 92 industrial centers from 1917–19 in order to provide appropriate weighting patterns for the new, more comprehensive index that would become the CPI. Periodic collection of prices was started and, in 1919, the Bureau of Labor Statistics (BLS) began publication of separate consumer price 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.

Since its inception, the CPI has been comprehensively revised on several occasions to implement the following: updated samples and weights, expanded coverage, and enhanced methodologies. The improvements introduced over the years have reflected not only the Bureau’s own experience and research, but also the criticisms and investigations of outside researchers.

Timeline

Key developments

1913: First year official retail price data are available

- In 1919 data were estimated back to 1913

1919: Began publication of separate indexes for 32 cities

- Collected prices in central cities periodically
- BLS used a 1917–19 study on family expenditures in 92 industrial centers to develop weights reflecting the relative importance of goods and services purchased by consumers.
- Collected prices for major groups: food, clothing, rent, fuels, house furnishings, and miscellaneous
- Limited pricing to items selected in advance to represent their categories

1921: Began regular publication of a national index, the U.S. city average, based on an unweighted average of the city indexes

- Estimated the U.S. city average back to 1913, using food prices only

1940: The first comprehensive revision of the CPI

- Used weights based on a 1934–36 study of consumer expenditures
- Collected prices in the 34 largest U.S. cities
- Implemented a weighted average of cities for the U.S. city average indexes

1942–43: Improvements made during World War II

- Discontinued the pricing of unavailable items, such as new cars and household appliances
- Increased the weight of other items, including automobile repair and public transportation

1950: Weight updates and new items added

- Adjusted weights in seven cities, using a 1947–49 survey of consumer expenditures
- Adjusted weights based on the 1950 Census
- Adjusted rent index to remove “new unit bias” caused by rent control
- Added new items to the list of covered items, including frozen foods and televisions

1953: The second comprehensive revision

- Used weights based on a 1950 survey of consumer expenditures conducted in central cities and attached urbanized areas
- Added a sample of medium and small cities
- Updated the list of items that the index covered, adding restaurant meals
- Expanded the sample of rental units and added homeownership costs and improved pricing and calculation methods

1964: The third comprehensive revision

- Based weights on a 1960–61 survey of consumer expenditures in metropolitan areas
- Added single-person households to the target population: urban wage earner and clerical worker households
- Extended price collection of goods and services to the suburbs of sampled metropolitan areas
- Updated the sample of cities, goods and services, and retail stores and service establishments

1966: Made quality adjustments for new vehicles at model changeover
1967: Improved treatment of seasonal items
1978: The fourth comprehensive revision

- Added a new Consumer Price Index: the CPI for All Urban Consumers, or the CPI-U
- Changed the name of the older CPI to the CPI for Urban Wage Earners and Clerical Workers, or the CPI-W
- Based weights on a 1972–73 survey of consumer expenditures and the 1970 census
- Expanded the sample to 85 areas (primary sampling units)

- Increased minimum frequency for obtaining the prices of goods and services (known as “pricing”) from quarterly to bi-monthly
- Implemented monthly pricing in the five largest areas
- Introduced probability sampling methods at all stages of CPI sampling
- Introduced checklists that define each category of spending to clarify what is included or excluded from an item (See appendix 2.)
- Developed estimates of the CPI’s sampling error and optimal sample allocation to minimize that error
- Began systematic replacement of retail outlets and their item samples between major revisions
    - Implemented a point-of-purchase survey (POPS)
    - Selected retail outlets with probability proportional to consumer spending therein
- Eliminated reliance on outdated secondary-source sampling frames
- Began rotating outlet samples every 5 years
- Began rotating one-fifth of the CPI pricing areas each year

1987: The fifth comprehensive revision

- Based weights on the 1982–84 Consumer Expenditure Survey and the 1980 Census
- Updated samples of items, outlets, and areas
- Began rotating item samples every 5 years
- Redesigned the CPI housing survey
- Improved sampling, data collection, data processing, and statistical estimation methods
- Initiated more efficient sample design and sample allocation
- Introduced techniques to make CPI production and calculation more efficient
- Introduced rental equivalence concept in January 1983 for the CPI-U and in January 1985 for the CPI-W (Additional information on rental equivalence is available in a factsheet.)

1988: Improved housing estimator to account for aging of the sample housing units
1992: Improved the handling of new models of vehicles and other goods
1994: Improved seasonal adjustment methods
1995: Implemented new sample procedures to prevent overweighting items whose prices are likely to rise
1997: Changes made to hospital services

- Initiated a single hospital services item stratum with a treatment-oriented item definition
- Discontinued pricing of the inputs to hospital services

1998: The sixth comprehensive revision

- Based weights on the 1993–95 Consumer Expenditure Survey and the 1990 census
- Updated geographic and housing samples
- Extensively revised item classification system
- Implemented new housing index estimation system
- Used computer-assisted data collection

- Added the Telephone Point-of-Purchase Survey (TPOPS) which allows rotation of outlet and item samples by item category and geographic area, rather than by area alone (this survey was previously conducted in person)

1999: Developed the Consumer Price Index research series

- Developed the Consumer Price Index Research Series, which is an estimate of the CPI for all Urban Consumers (CPI-U) from 1978 to present that incorporates most of the improvements made over that time span into the entire series

1999: Initiated a new housing survey based on the 1990 census

- Initiated a new housing survey based on the 1990 census that estimated price change for owners' equivalent rent directly from rents (an estimate of the implicit rent owner occupants would have to pay if they were renting their homes)

1999: Began using a geometric mean formula for most basic indexes

- Began using a geometric mean formula for most basic indexes that mitigates lower level substitution bias and reflects shifts in consumer spending with item categories as relative price change

1999–2000: Expanded the use of hedonic regression in quality adjustment
2002: Addition of the Chained Consumer Price Index for All Urban Consumers (C-CPI-U) and other changes

- Implemented biennial weight updates
- Implemented 4-year outlet rotation to replace 5-year scheme
- Added the C-CPI-U in August
    - Uses more advanced "superlative" index formula (the Törnqvist formula)
    - Corrects upper-level substitution bias

2003: Directed replacement of sample items in the personal computer and other categories, to keep samples current
2004: Expanded collection of price data to all business days of the month

- Expanded collection of price data to all business days of the month (Before 2004, prices were collected during the first 18 business days of the month for the first 10 months of the year and the first 15 business days for November and December.)

2007: Began publishing indexes to three decimal places (January)
2008: Added CPI-E and published it back to 1982

- Congress mandated experimental consumer price index for older people (62 years of age and older)
- CPI-E has limitations related to the expenditure weights, outlets sampled, items priced, and the prices collected

2010: Began a three-stage effort to improve the housing survey

- First and second stages used the 2000 census
- First stage was a 4-year sample augmentation with a goal of adding 16,000 units, mainly in neighborhoods with seriously depleted renter samples; this data was first used in the rent indexes for July 2010
- Second stage was a sample replacement meant to replace the rental units introduced in 1999. The November 2012 CPI was the first that used a new sample from this stage and the May 2016 CPI was the first in which the housing survey sample was drawn entirely from the 2000 census.
- The third stage was a regular replacement commencing in 2016 that will end in 2022. It will replace the 2000 census-based sample with one based on the American Community Survey using 2010 Decennial census geography. This stage will continue into the future and, for the first time, the CPI housing survey will have a process that keeps its sample continuously updated.

2018: Geographic revision

- Introduced a new geographic sample based on the 2010 Census
- Changed the publication frequency for several local area indexes
- Established new local area and aggregate indexes
- Introduced census division-level index data

2019: Changes to the CPI establishment frame

- Replaced Telephone Point-of-Purchase Survey (TPOPS) as source of retail establishment frame with data from the Consumer Expenditure Surveys (CE)
- Eliminates redundancies and inefficiencies in survey operations and reduces household burden
- Use of Quarterly Census of Employment and Wages business registry to refine the location and address data from the CE

NOTES

[1] Rippy, Darren. "The first hundred years of the Consumer Price Index: a methodological and political history." *Monthly Labor Review,* April 2014

[2] This discussion follows *The Consumer Price Index: history and techniques*, Bulletin No. 1517 (U.S. Bureau of Labor Statistics, 1966), p. 1-2.

Archives

- February 14, 2018

Last Modified Date: November 24, 2020

More Information

Additional information on the Consumer Price Index (CPI) is available on the CPI website: www.bls.gov/cpi.

CPI data are available on the BLS website:

- Databases: www.bls.gov/cpi/data.htm
- Lastest news releases: https://www.bls.gov/cpi/news.htm

For some of our common questions and answers see www.bls.gov/cpi/questions-and-answers.htm.

Technical references Aizcorbe, Ana M., and Patrick C. Jackman. "The commodity substitution effect in CPI data, 1982–91." *Monthly Labor Review*, December 1993, pp. 25–33.

Blair, Caitlin. "Constructing a PCE-Weighted Consumer Price Index." National Bureau of Economic Research, NBER Working Paper No. 19582. October 2013.

Brown, Craig and Anya Stockburger. "Item replacement and quality change in apparel price indexes." *Monthly Labor Review,* December 2006, pp. 35-45

Cage, Robert. "New methodology for selecting outlet samples." *Monthly Labor Review*, December 1996, pp. 49–61.

The Conference Board. "Measuring Prices in a Dynamic Economy: Re-examining the CPI." New York: The Conference Board, 1999.

Dalton, Kenneth V., John S. Greenlees, and Kenneth J. Stewart. "Incorporating a geometric mean formula into the CPI." *Monthly Labor Review*, October 1998, pp. 3–7.

Diewert, W. E., and A. O. Nakamura, eds. *Essays in index number theory,* vol. 1. Amsterdam: North-Holland Publishing Co., 1993.

Final report of the advisory commission to study the Consumer Price Index. U.S. Senate, Committee on Finance, 104th Cong., 2d sess., 1996.3

Fixler, Dennis, Charles Fortuna, John Greenlees, and Walter Lane. "The use of hedonic regressions to handle quality change: the experience of the U.S. CPI." Paper presented at the Fifth Meeting of the International Working Group on Price Indices, Reykjavik, Iceland, August, 26,1999.

Gillingham, Robert F. "A conceptual framework for the Consumer Price Index." *Proceedings of the American Statistical Association*, Business and Economics Section. Alexandria, VA: American Statistical Association, 1974, pp. 46–52.

Gillingham, Robert F. "Measuring the cost of shelter for homeowners: theoretical and empirical considerations." *The Review of Economics and Statistics*, vol. 65, no. 2, May 1983, pp. 254–265.

Greenlees, John S., and Robert B. McClelland. "Addressing misconceptions about the Consumer Price Index." *Monthly Labor Review*, August 2008, pp.3–17.

Hulten, Charles R. "Quality Change in the CPI," Federal Reserve Bank of St. Louis Review, May 2001, pp. 87–111.

Johnson, David S, Stephen B. Reed and Kenneth J. Stewart. "Price Measurement in the United States: A decade after the Boskin Report" *Monthly Labor Review*, May 2006 pp. 10-19.

Klick, Joshua. "Improving initial estimates of the Chained Consumer Price Index." *Monthly Labor Review*, February 2016.

Konny, Crystal G., Brendan K. Williams, and David M. Friedman. "Big Data in the U.S. Consumer Price Index: Experiences & Plans." Paper presented at the National Bureau of Economic Research conference, "Big Data for 21st Century Economic Statistics," Washington, D.C., March 15-16, 2019.

Liegey, Paul. "Hedonic quality adjustments in the U.S. CPI: a statistical agency perspective." Paper presented at Center for European Economic Research Conference, "Price indices and the measurement of quality change," University of Mannheim, Germany, April 25–26, 2002.

Moulton, Brent R. "Basic components of the CPI: estimation of price changes." *Monthly Labor Review*, December 1993, pp. 13–24.

Moulton, Brent R., and Karin E. Moses. "Addressing the quality exchange issue in the Consumer Price Index." *Brookings Papers on Economic Activity*, Washington: The Brookings Institution, 1997.

Paben, Steven P., William H. Johnson, and John F. Schilp. "The 2018 revision of the Consumer Price Index geographic sample." *Monthly Labor Review, October 2016.*

Pollak, Robert A. The theory of the cost-of-living index. Oxford, U.K: Oxford University Press, 1989.

Ptacek, Frank. "Updating the rent sample for the CPI Housing Survey." *Monthly Labor Review,* August 2013.

Reed, Stephen. "One hundred years of price change: the Consumer Price Index and the American inflation experience." *Monthly Labor Review,* April 2014.

Reed, Stephen, and Stewart, Kenneth J. "Why does BLS provide both the CPI-W and CPI-U?." *Beyond the Numbers,* February 2014.

Rippy, Darren. "The first hundred years of the Consumer Price Index: a methodological and political history." *Monthly Labor Review,* April 2014.

Charles L. Schultze, and Christopher Mackie, eds. *At What Price? Conceptualizing and measuring cost-of-living and price indexes*. National Research Council Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes, Committee on National Statistics, Commission on Behavioral and Social Sciences and Education. Washington: National Academies Press, 2002.

Shoemaker, Owen J., and William H. Johnson. "Estimation of variance components for the U.S. Consumer Price Index." *Proceedings of the American Statistical Association*, Government Statistics and Social Statistics sections. Alexandria, VA: American Statistical Association, 1999, pp. 298–303.

Stigler, George, ed. "The price statistics of the federal government." In *Report to the Office of Statistical Standards, Bureau of the Budget*. New York: National Bureau of Economic Research, 1961.

U.S. Bureau of Labor Statistics, "Measurement Issues in the Consumer Price Index," *Statistical Journal of the United Nations* ECE, 15 (1998), pp.1–36.

Williams, Eilliot. "The effects of rounding on the Consumer Price Index." *Monthly Labor Review,* October 2006, pp. 80-89.

Williams, Janet L., Eugene F. Brown, and Gary R. Zion. "The challenge of redesigning the Consumer Price Index area sample." Proceedings of the American Statistical Association, Survey Research Methods Section. Alexandria, VA: American Statistical Association, 2003, pp. 200–205.

Appendix 1. CPI Geographic sample, 2018

Appendix 2. Content of CPI entry level items

Appendix 3. Consumer Expenditure sample rotation categories

Appendix 4. Non-Consumer Expenditure sample rotation categories

Appendix 5. Consumer Expenditure UCC to CPI ELI concordance

Appendix 6. Items for which average price data is published

Appendix 7. CPI items by publication level

Have more questions? For ways to contact us see www.bls.gov/cpi/contact.htm

Last Modified Date: November 24, 2020