**CPI-U Calculation Process**

The CPI is built in two stages (shown in Figure 1).

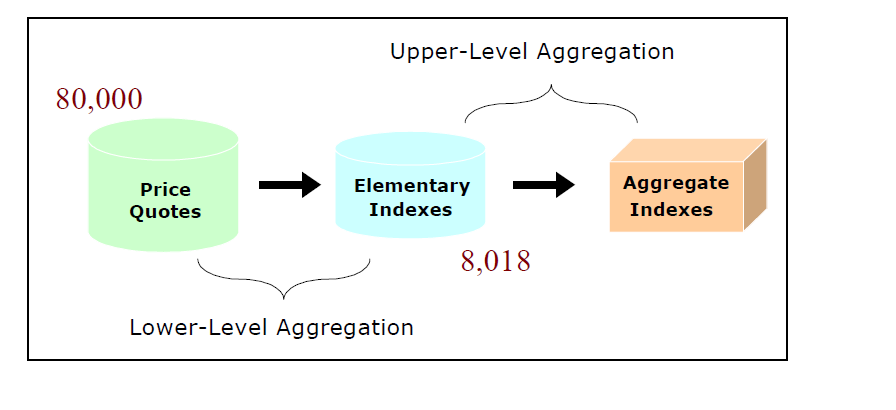


Figure 1 2-Stage Construction of the CPI

**1.1 Lower-level aggregation (Elementary-level index)**

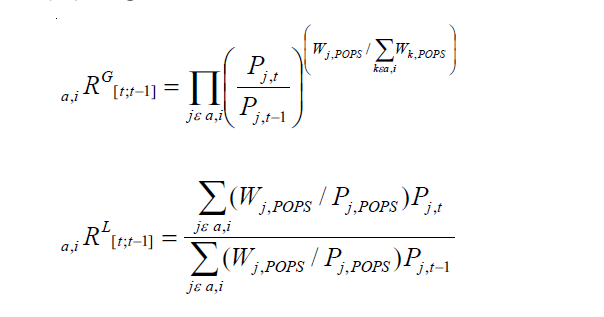
The first stage is often referred to as “lower-level aggregation” or “elementary-level aggregation” as it involves averaging the most fundamental component of the index - observed price change for specifically defined consumer goods, services, and products. In this stage, price changes for roughly 80,000 specific items per month are averaged to yield 8,018 estimates of aggregate price change.

In the CPI, the urban portion of the United States is divided into 38 geographic areas called *index areas*, and the set of all goods and services purchased by consumers is divided into 211 categories called *item strata.* This results in 8,018 (38 x 211) item-area combinations. An elementary-level index is computed for each combination of an item stratum and index area.

The CPI item structure has four levels of classification. The 8 major groups are made up of 70 expenditure classes (ECs), which in turn are divided into 211 item strata. Major groups and ECs do not figure directly in CPI sample selection, although ECs are used in smoothing item stratum expenditure estimates during composite estimation. Within each item stratum, one or more substrata, called entry-level items (ELIs), are defined. There are a total of 305 ELIs, which are the ultimate sampling units for items as selected by the BLS national office. They represent the level of item definition from which data collectors begin item sampling within each sample outlet. Each ELI has a corresponding sampling frame of outlets that sell the ELI. In each there are several unique items. A single selection of a unique item is referred to as a quote. An example of a unique item for the ELI Cookies is a 1-lb. bag of chewy-style chocolate-chip cookies with walnuts, of a particular brand name.

For example, the prices of approximately ten different brands and styles of watches at various locations in Chicago are observed each month, compared to the prices observed in the previous month, and averaged together to produce an index of price change for watches in Chicago. Watches (ITEM=AG01) is one of 211 elementary items, and Chicago (AREA=A207) is one of 38 elementary areas in the current CPI market basket structure. The Chicago-watch index is one of the 8,018 (211 items x 38 areas) elementary indexes produced in the first stage of CPI construction.

According to the BLS’s handbook for Consumer Price Index, The weight for the lower-level index is not based on expenditure survey data. The weights for the first stage come from the sampling frame for the category[[1]](#footnote-1). Since January 1999, most item strata have used an expenditure-share-weighted geometric average, *a,tRG[t,t-1]*. The other strata use the Laspeyres formula average, *a,tRL[t,t-1]*, which all strata used prior to 1999. The Laspeyres is a base-period quantity-weighted arithmetic average. Every month the C&S system computes price relatives for each area-item combination (*a,i*) using these formulas:



where

*a,tRG[t,t-1]*  and *a,tRL[t,t-1],* are, respectively, the geometric and Laspeyres price relatives for area-item combination, *a,i*, from the previous period, *t-1* (either 1 month or 2 months ago), to the current month, *t*;

*Pj,t* is the price of the *j*th observed item in month *t* for area-item combination *a,i*;

*Pj,t-1* is the price of *the same item* in time *t-1*;

*Pj,POPS* is an estimate of the item *j*’s price in the sampling period when its POPS was conducted;

*Wj,POPS* is item j’s weight in the POPS, which is defined in detail below.

Fuxa (2010) provides detailed information about how the Consumer Expenditure data are used in calculating elementary-level index.

An index *I0,t*, reflecting price change from month 0 to the current month *t*, is calculated at the item stratum and index-area level by multiplying the index for the previous month by a price relative that measures the price change between the previous month and the current month:

where: *I0,(t-1)* = index from 0 to (t-1); *R(t-1),t*  = price relative from (t-1) to t.

Next equation shows the geometric mean that is the target price relative for all but a few item strata. For the sake of simplicity, it is shown at the level of a PSU-replicate intersection.

where: *N*= number of ELIs in the item stratum; *NE*= number of outlets in the frame corresponding to ELI *E*; *NE,O* = number of unique items in outlet *O* for ELI *E*; *(Pt/Pt-1)E,O,U* = the ratio of the price in month *t* to the price in month *t*-1 for unique item *U* in outlet *O* for ELI *E*; (*va*)*E,O,U* = consumer expenditure in month *a* for unique item *U* in outlet *O* for ELI *E*, where month *a* is associated with the timing of sample rotation.

**2. Estimation of upper-level price change**

In the second stage, the elementary indexes are averaged together to yield various aggregate indexes and ultimately the All-Items, U.S. City Average index of price change.

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

Following sections illustrate how to estimate up-level price change. (The whole process is copied from page 33-36 of Chapter 17 Consumer Price Index by BLS[[2]](#footnote-2), this document illustrate how Consumer Expenditure are used in detail.)

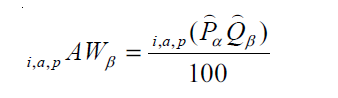
**2.1 Input elementary price indexes**

BLS field economists do not price the samples of quotes in all 8,018 elementary item-area cells on a monthly basis. Approximately 40 percent of the elementary item-area combinations are priced on a bi-monthly schedule, with half of these priced every even month and the other half priced every odd month. Months for which prices are not collected are called “off-cycle” months.

For CPI-U, off-cycle bimonthly elementary indexes are set equal to the previous-month index value.

**2.2 Input elementary expenditure weights**

To aggregate elementary indexes into published indexes, an aggregation weight for each elementary item-area combination is required. The function of the aggregation weight is to assign each elementary 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 211 elementary items in the 38 elementary areas comprising the CPI sample, for a specified time period. In the CPI-U and CPI-W, aggregation weights (AW) are defined as:



where i,a,pα is the estimated price of item (i) purchased in area (a) by population (p) in period (α), i,a,pα is the estimated quantity of item (i) purchased in area (a) by population (p) in period (β).

Period (α) is the base period of the corresponding elementary item-area index. For example, the “Sports equipment” (ITEM = RC02) in Seattle (AREA = A423) index has a base period of α = June 1985. CPI elementary indexes have varying base periods. Most published indexes have an index base period of α = 1982-84.

Time period (β) corresponds to the reference period of the expenditures used to derive the implicit quantity weights needed for Laspeyres aggregation. As of 2005, the CPI-U and CPI-W had an expenditure reference period of β = 2001-02. Historically, the CPI expenditure reference period has been updated approximately every 10 years. In 2002, BLS instituted a biennial rotation schedule for updating the expenditure reference period. Effective with the January 2004 index, the expenditure reference period changed from β = 1999-2000 to β = 2001-02; effective with the January 2006 index, it was updated again to 2003-04; 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 elementary index, but not the implicit price component (P) of the aggregation weight (AW) of each elementary index.

Aggregation weights for the CPI-U and CPI-W are derived from estimates of household expenditures collected in the CE. Despite an increase in the CE sample size in 1999, expenditure estimates at the elementary 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 elementary 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. Second, elementary item-area expenditures are averaged, or composite estimated, with item-regional expenditures. This has the effect of lowering the variance of each elementary 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 following equations.

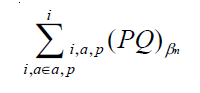
The estimated expenditure i,a,p()βn for item (i) in area (a) for population (p) in reference period (β) is derived from a weighted average of **the item’s relative importance in the elementary area** (a) and **its relative importance in its corresponding region-size classification (m),** for each year encompassing reference period (β). The weight (δ) assigned to the region-size class (m) and the weight (1-δ) assigned to the elementary area (a) are a function of the variance in each area and the covariance of each measure.47 The resulting average share ( is then multiplied by the sum of all expenditures in the elementary 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 (βn) have a lower bound of one cent ($0.01). The raked item expenditures in each year of reference period (β) are then averaged to obtain the estimated expenditure in (β). Finally, the estimated expenditure is adjusted by the corresponding item-area index to obtain the aggregation weight: an expenditure value with an implicit price of time period (α) and implicit quantity of time period (β).

The process of estimating CPI-U Elementary Aggregation Weights is as follows:

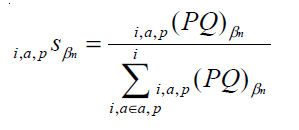
1. Expenditure on item (*i*) in area (*a*) by population (*p*) in year (*βn*)

*i,a,p(PQ)βn*

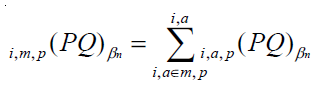
2. Total expenditures in area (*a*) by population (*p*) in year (*βn*)



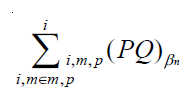
3. Share of total expenditures for item (*i*) in area (*a*) for population (*p*) in year (*βn*)



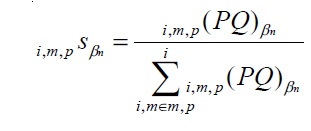
4. Expenditure on item (*i*) in major area (*m*) by population (*p*) in year (*βn*)



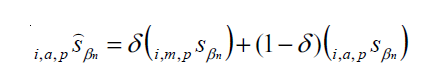
5. Total expenditures in major area (*m*) by population (*p*) in year (β*n*)



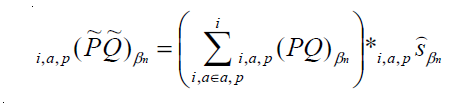
6. Share of total expenditures for item (*i*) in area (*m*) for population (*p*) in year (*n*)



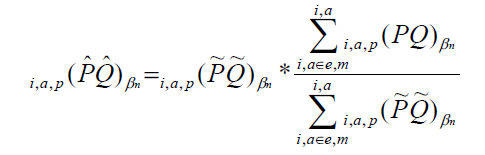
7. Composite-estimated share of total expenditures for item (*i*) in area (*a*) for population (*p*) in year (*n*)



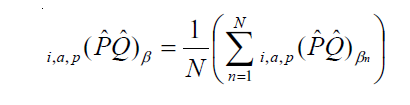
8. Estimated expenditure on item (*i*) in area (*a*) by population (*p*) in year (*n*)



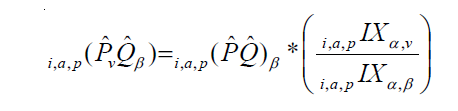
9. Raked expenditure on item (*i*) in area (*a*) by population (*p*) in year (*n*)



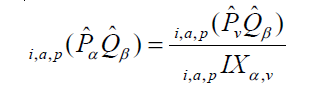
10. Estimated expenditure in expenditure reference period ()



11. Cost weight in pivot month (*v*)



12. Aggregation weight



where

*p* = population (urban or urban wage earner)

*a* = CPI elementary area

*i* = CPI elementary item

*e* = expenditure class

*m* = One of eight CPI major areas, defined by region-and-city-size classification. Regions

are Northeast, Midwest, South, and West; city-size types are self-representing and

non-selfrepresenting.

*P* = price

*Q* = quantity

*N* = number of years in the CPI-U expenditure reference period (NOTE: currently N=2.)

*βn* = year belonging to expenditure reference period β (NOTE: n=1 is 1999 and n=2 is 2000

in the current CPI-U expenditure reference period.)

δ = weight assigned to major area (*m*), where 0 < δ < 1

α = lower-level index base period

*v* = year and month, usually December, prior to the month when expenditure weights from

reference period β are first used in the CPI

*i,a,pSβn* = estimated expenditures (*PQ*) for item (*i*) in area (*a*) for population (*p*) as a percent of

total CPI expenditures in area (*a*) in period β*n*

i,a *IX*α,β = lower-level index of price change from index base period (α) to expenditure reference

period (β) for item (*i*) in area (*a*)

i,a *IX*α,*v* = lower-level index of price change from index base period (α) to pivot-month *(v)* for

item (*i*) in area (*a*)

According to Schmidt (1995), the weight is an expenditure-population weight. To calculate the CPI an estimate of the expenditure pattern is needed to define and weight the market basket of goods and services for which the index is computed. Each expenditure-population weight is the product of estimates of mean expenditures per consumer unit derived from the 1982-84 Consumer Expenditure Survey and estimates of the number of consumer units obtained from the 1980 census files. Mean expenditures are calculated using preliminary expenditures and their relative importance, composite estimated and raked to minimize the average mean square error of the relative importance of the total weight for the index. (Expenditure-population weight)

Each consumer unit in the survey is assigned a weight that is the number of similar consumer units in the U.S. civilian noninstitutional population the sampled consumer unit represents. Using these weights, the average expenditure per consumer unit on a particular item category is estimated with the formula[[3]](#footnote-3):

where

= the average expenditure per consumer unit on the item category,

*yi* = the expenditure made by the *ith* consumer unit on the item category,

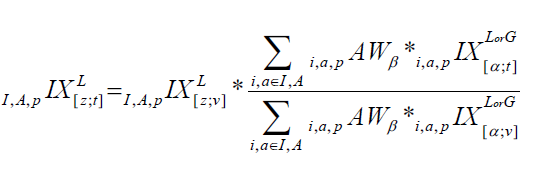
*wi* = the weight of the *ith* consumer unit in the sample, and

*S* = the sample of consumer units that participated in the survey.

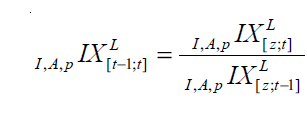
2.3 ***Aggregation formula.***

The Laspeyres price index is used to aggregate elementary 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 elementary item-area index. These quantity weights remain fixed for a 2-year period, and then are replaced each January in each even 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 index divided by the index value in the index base period. (See the equation for long-term price change below.) Month-to-month price change is then calculated as a ratio of the long-term monthly indexes. The CPI-U upper-level aggregation formula is as follows:

Long-term price change



Month-to-month price change



where

*A* = all elementary areas (“U.S. city average”)

*a*  = CPI elementary area

*p* = population (the C-CPI-U is calculated for the U-population only.)

*i* = CPI elementary item

*I* = all elementary items (“all-items”)

*t* = month

*z* = base period of the aggregate index (NOTE: the U.S. city average—all-

items CPI index has a base period of z=1982-84.)

α = base period of the elementary index (*i*) in area (*a*)

*v* = year and month, usually December, prior to the month when expenditure

weights from reference period (β) are first used in the CPI

*i,a,p IX[α;t]* = lower-level index of price change from period (α) to month (*t*) for item

(*i*) in area (*a*)

*i,a,p IX[α;v]* = lower-level index of price change from period (α) to pivot-month (*v*) for

item (*i*) in area (*a*)

*i,a,pAWβ* = aggregation weight from reference period (β) for item (*i*) in area (*a*)

*I,A,pIXL[z;v]* = aggregate-level CPI-U index of price change from period (*z*) to pivot-

month (*v*) for aggregate item (*I*) in aggregate area (*A*) for population (*p*)

**3. How does BLS decide certain quote are included in maker basket or not?**

In frequent asked questions for CPI, BLS answers how to determine the market basket[[4]](#footnote-4). The CPI market basket is developed from detailed expenditure information provided by families and individuals on what they actually bought. For the current CPI, this information was collected from the Consumer Expenditure Surveys for 2011 and 2012. In each of those years, about 7,000 families from around the country provided information each quarter on their spending habits in the interview survey. To collect information on frequently purchased items, such as food and personal care products, another 7,000 families in each of these years kept diaries listing everything they bought during a 2-week period.

Over the 2 year period, then, expenditure information came from approximately 28,000 weekly diaries and 60,000 quarterly interviews used to determine the importance, or weight, of the more than 200 item categories in the CPI index structure.

In frequent asked questions for CPI, BLS answers what goods and services the CPI cover[[5]](#footnote-5). The CPI represents **all goods and services purchased for consumption** by the reference population (U or W) BLS has classified all expenditure items into more than 200 categories, arranged into eight major groups.

Also included within these major groups are various government-charged user fees, such as water and sewerage charges, auto registration fees, and vehicle tolls. In addition, the CPI includes taxes (such as sales and excise taxes) that are directly associated with the prices of specific goods and services. However, the CPI excludes taxes (such as income and Social Security taxes) not directly associated with the purchase of consumer goods and services.

The CPI does not include investment items, such as stocks, bonds, real estate, and life insurance. (These items relate to savings and not to day-to-day consumption expenses.)

For each of the more than 200 item categories, using scientific statistical procedures, the Bureau has chosen samples of several hundred specific items within selected business establishments frequented by consumers to represent the thousands of varieties available in the marketplace. For example, in a given supermarket, the Bureau may choose a plastic bag of golden delicious apples, U.S. extra fancy grade, weighing 4.4 pounds to represent the Apples category.

To enable the CPI to reflect changes in the marketplace, new item and outlet samples are selected each year, on a rotating basis, for approximately 25 percent of the item strata in each PSU. Each year, four regional item universes are tabulated from the 2 most recent years of CE data. Independent samples of ELIs are selected from the corresponding regional item universe for each item stratum PSU replicate scheduled for rotation that year. Within each sample PSU-replicate, each item sample is based on a systematic probability-proportional-to-size (PPS) sampling procedure, in which each ELI has a probability of selection proportional to the CPI-U population expenditures for the region for the ELI within its stratum.

**4. How does BLS get the relative importance of items?**

According to Schmidt (1995), the relative importance based on CE data refers to the ratio of expenditure of an item or a group of items to the total expenditures for all items. The patterns of expenditures at major group levels are quite similar over these time periods.

According to Mason and Butler (1987), the expenditure weight for each item stratum is an estimate of total expenditure by the index population for that item. It is calculated as the product of estimates of mean expenditures of consumer units and the number of consumer units. Mean expenditures are derived from 1982-84 Consumer Expenditure Survey data, and estimates of the number of consumer units are obtained from a special tabulation of the 1980 census files.

Mason and Bulter (1987) defines relative importance as the share that the base-period expenditure multiplied by the price relative for a particular item stratum is of the sum of all base-period expenditures multiplied by their price relatives.

Where

*Pti* is the price of item I in the comparison period t;

*Poi* is the price of item I in the base period;

*Eoi* is the expenditure for item i in the base period

**5. Purpose of CE survey**

The objectives of the CE remain the same: to provide the basis for revising weights and associated pricing samples for the CPI and to meet the need for timely and detailed information on the spending patterns of different types of families. (Weight and associated pricing samples)

Results of the CE are used to select new “market baskets” of goods and services for the index, to determine the relative importance of components, and to derive cost weights for the baskets.

The BLS describe how CE data are used in the CPI[[6]](#footnote-6).

1. The BLS combines the spending information from respondents across the country to see how much is spent on each type of item.
2. All reported expenses are used to estimate how much urban households spend on each item.
3. These estimates are used to construct the market basket which contains a representative sample of expenses.
4. The BLS conducts another survey to find out where consumers purchased items in the market basket.
5. BLS data collectors visit housing units and a sample of the identified stores to obtain current price information on about 80,000 items each month across the country.
6. The BLS combines the information about the items purchased, the expenditures on these items, and their current prices to calculate the CPI.

**Reference**

Fuxa, Mary Lee. 2010. “Sampling and Weighting of Commodity and Service Units for the Elementary Level of Computation of the US Consumer Price Index.” *Proceeding Business and Economic Statistics Section, American Statistical Association*, 5014–25.

Mason, Charles, and Clifford Butler. 1987. “New Basket of Goods and Services Being Priced in Revised CPI.” *Monthly Labor Review*, 3–22.

Schmidt, Mary Lynn, and Bureau Of Labor Statistics. 1995. *Comparing Market Basket Changes and the CPI*.

1. Chapter 17 The Consumer Price Index by BLS, page 21, available at <http://www.bls.gov/opub/hom/pdf/homch17.pdf> [↑](#footnote-ref-1)
2. Chapter 17 Consumer Price Index by BLS, available at <http://www.bls.gov/opub/hom/pdf/homch17.pdf> [↑](#footnote-ref-2)
3. Chapter 16 Consumer expenditures and income, page 6, available at <http://www.bls.gov/opub/hom/pdf/homch16.pdf> [↑](#footnote-ref-3)
4. BLS Frequent Asked Questions CPI, how is the CPI market basket determined? Available at <http://stats.bls.gov/cpi/cpifaq.htm#Question_6> [↑](#footnote-ref-4)
5. BLS Frequent Asked Questions CPI, what goods and services does the CPI cover? Available at <http://stats.bls.gov/cpi/cpifaq.htm#Question_7> [↑](#footnote-ref-5)
6. The CE and the Consumer Price Index, available at <http://www.bls.gov/respondents/cex/ceandcpi.htm> [↑](#footnote-ref-6)