

Overview of the Consumer Price Index (CPI) Program

- Goal of the CPI
- History and Milestones
- Scope / Coverage
- Classification Systems
- Weighting
- Sampling
- Data Collection
- Estimation
- Publication
- Challenges and Future Improvements



Goal of the CPI

- The goal of the CPI is to approximate a cost of living index.
- Cost of living is a theoretical concept. The CPI seeks to measure the change in the cost of living by measuring the change in prices that consumers pay for a market basket of goods and services.

Brief history of the CPI

- BLS price and expenditure data first collected in 1888. BLS published food price measures as early as 1890
- During the World War I era, a consensus emerges that existing price data, predominantly food price measures, was inadequate as a cost-of-living measure
- More comprehensive price collection in 92 cities began in 1917, with publication of data in 1919
- Regular publication began in 1921, with estimates created back to 1913

CPI milestones

- 1940: First comprehensive revision of the CPI
- 1973: Amendment to Social Security Act ties Social Security payments to CPI. First applied in 1975
- 1978: Fourth major revision includes creation of CPI-U; older CPI series continues as CPI-W
- 1981: Economic Recovery Tax Act indexes many aspects of the tax code to the CPI

CPI milestones

- 1996: Boskin Commission report focuses attention on the CPI
 - ▶ Asserts upward bias in the CPI based mostly on substitution bias, new goods bias, and quality change bias
- 1998-2000: CPI expands use of hedonic adjustment and other direct quality adjustment methods to better measure quality change
- 1999: CPI moves to Geometric means formula for most basic indexes
 - ▶ Addresses lower level substitution bias

CPI milestones

- 2002: Chained Consumer Price Index (C-CPI-U) is introduced
 - ▶ Published backed to 1999
 - ▶ Addresses upper level substitution bias
- 2002: CPI moves to biennial weight updates
 - ▶ Previously, weight updates had occurred only with major revisions (every 10 years or so)
 - ▶ Reduces new goods bias since newer items are incorporated more quickly into the CPI

Scope / Coverage

- The CPI reflects prices paid by consumers in urban areas of the U.S. for a market basket of goods and services.
 - ▶ “Urban” in this context is broadly defined and encompasses 89 percent of the population
- Any item purchased for consumption, goods or services, is eligible for pricing.

Classification systems

- In the CPI, the consumer market basket is a categorization of goods and services with corresponding weights. In the U.S. there are 211 categories of items (goods and services) that are aggregated into eight major groups:
 - Food And Beverages
 - Housing
 - Apparel
 - Transportation
 - Medical Care
 - Recreation
 - Education and Communication
 - Other Goods and Services

Classification systems

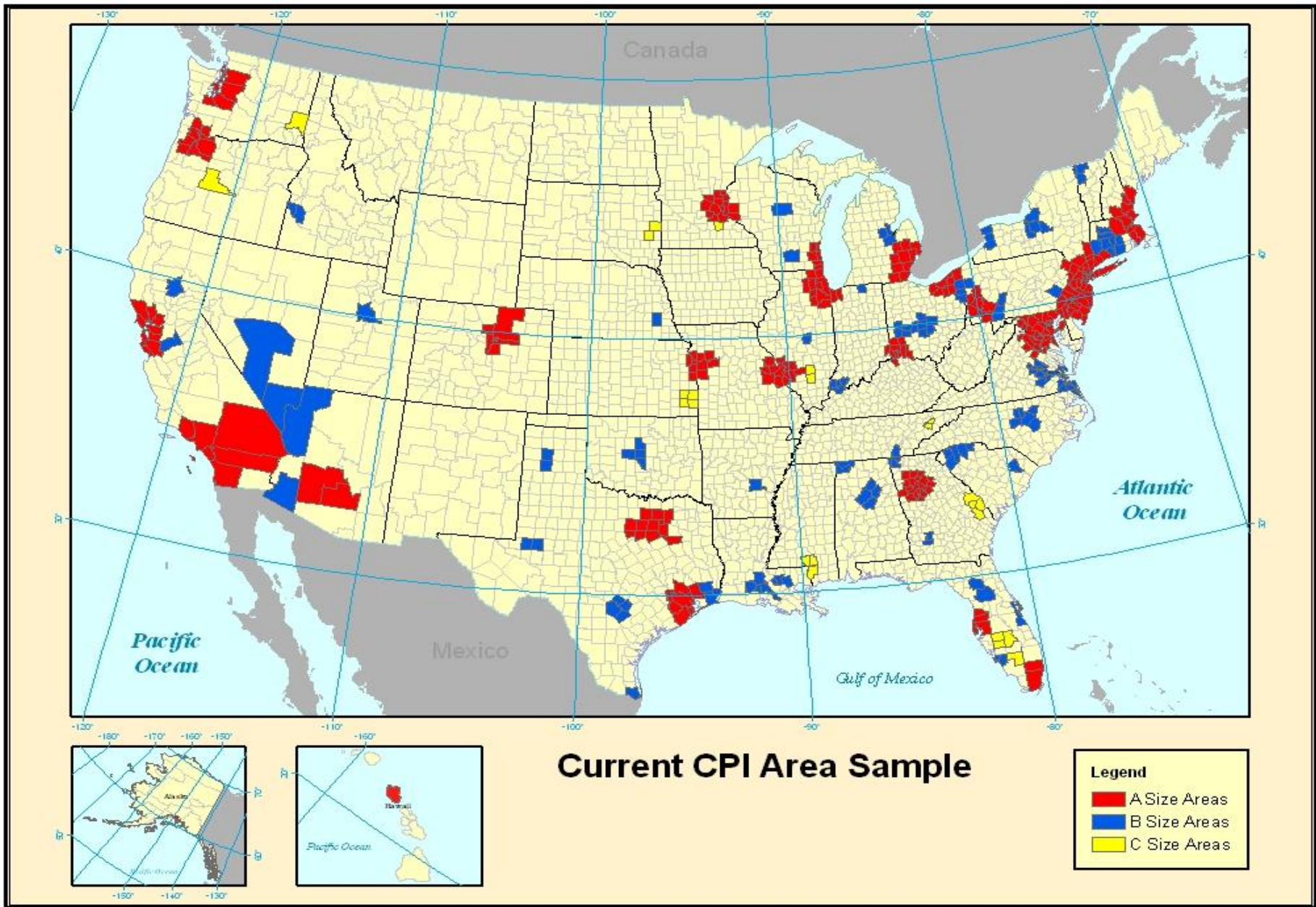
- Additionally, the CPI can be divided into food, energy, and all items less food and energy:
- All Items
 - ▶ Food
 - Food at home
 - Food away from home
 - ▶ Energy
 - Energy commodities
 - Energy Services
 - ▶ All items less food and energy

Sampling / Geography

- The CPI reflects prices paid by consumers in urban (broadly defined) areas of the U.S.
- Based on Census and OMB definitions, 87 geographic areas are selected to represent the urban population.
- These 87 areas are referred to as Primary Sampling Units (PSUs).

Sampling / Geography

- These 87 areas represent 38 distinct geographical units (31 large cities plus a combination of the remaining 56 smaller areas into seven geographical units).
 - ▶ The seven units include medium size cities in each region and small cities in all regions except the Northeast
- The following slide shows the current sample, based on the 1990 census
- Work is ongoing to bring in a new geographic sample



Data collection / Outlets

- The geographic sample and market basket form a matrix of component cells from which to build a CPI.
- Item categories x Geography Units = Cells
 - ▶ 211 item categories x 38 geographic units = 8018 cells
 - ▶ Basic indexes calculated for each cell

Data Collection / Outlets

- The Telephone Point-of-Purchase Survey of households (TPOPS) is used to create the frame of outlets for the collection of data in each of the 87 PSUs.
- BLS Field Representatives visit outlets and use Computer Assisted Data Collection (CADDC) to select and price items using probability sampling.
- Items are described completely in terms of price determining characteristics using CPI Checklists.

Data Collection / Outlets

- ▶ Sample rotation allows the sample of specific items in the CPI to stay up-to-date.
 - This allows, for example, the particular cell phones or automobiles in the CPI sample to be representative of what consumers are buying
- ▶ The outlet/item sample are replaced every four years, 1/8th of the sample every 6 months

Data Collection / Outlets

- Based on the TPOPs survey, the commodity product or service line(s) (e.g., apples) to be priced in each outlet is known in advance of visiting the outlet.
- During the initiation interview at an outlet, a process known as disaggregation is performed to determine the exact items to be priced within these product or service lines.

Data Collection / Outlets

- The exact items are chosen using statistical methods that give individual items a chance to be selected proportional to their sales for that particular product or service line at the selected outlet.
- CPI price data are collected throughout the entire month. The month is divided into three pricing periods, with field representatives required to collect data during each period.
- In most cities, prices on many items are collected bimonthly rather than monthly

Data Collection / Outlets

- The price sought in the CPI is the retail, transaction price paid by the consumer, including sales and excise taxes.
- Each month a field staff of around 400 part-time economic assistants and 100 full-time economists collect prices for over 83,000 individual items per month (around a million prices per year) based on personal visits to more than 23,000 outlets in 87 cities.

Data Collection / Housing

- One of the most important measures in the CPI is tracking the rate of inflation in housing services through a survey that collects data from around 5800 housing rental units per month, about 69,000 annually
- The rent sample serves two purposes:
 - ▶ First, it provides the data for measuring changes in shelter for consumers who rent (about one third of all consumers).
 - ▶ Second, the CPI rent sample provides the data for measuring changes in shelter costs for consumers who own their own homes (about two thirds of all consumers).

Data Collection / Housing

- This latter index is estimated based on the concept of rental equivalence or the market rent that would be charged for these owner occupied dwellings if they were rented.
- In the CPI, increases (or decreases) in the market value of a home are not viewed as the proper measure of housing inflation.
- Rather, we strive to measure the payments made by consumers in a given month to obtain the services of a home (aka shelter). We call this concept Owners' Equivalent Rent (OER).

Weighting

- Conceptually, the CPI seeks a market basket that is representative of what is actually purchased by consumers. We know bananas should be in the market basket, but how much weight should they have?
- The source of weighting for the CPI is the Consumer Expenditure Survey (CE).



Weighting

- The CE consists of two separate surveys –a quarterly interview and a diary:
 - ▶ The quarterly interview surveys (over 5 consecutive quarters) are used to ask consumers about their major purchases.
 - ▶ Additionally, a sample of consumers keep diaries of their purchases over two single two week period.

Weighting

- These surveys are used to create the expenditure weights that are used in constructing the CPI market basket.
- The expenditure weights are updated every two years.
- The weights used in the CPI in 2014 & 2015 are based on consumer expenditures in 2011-2012.

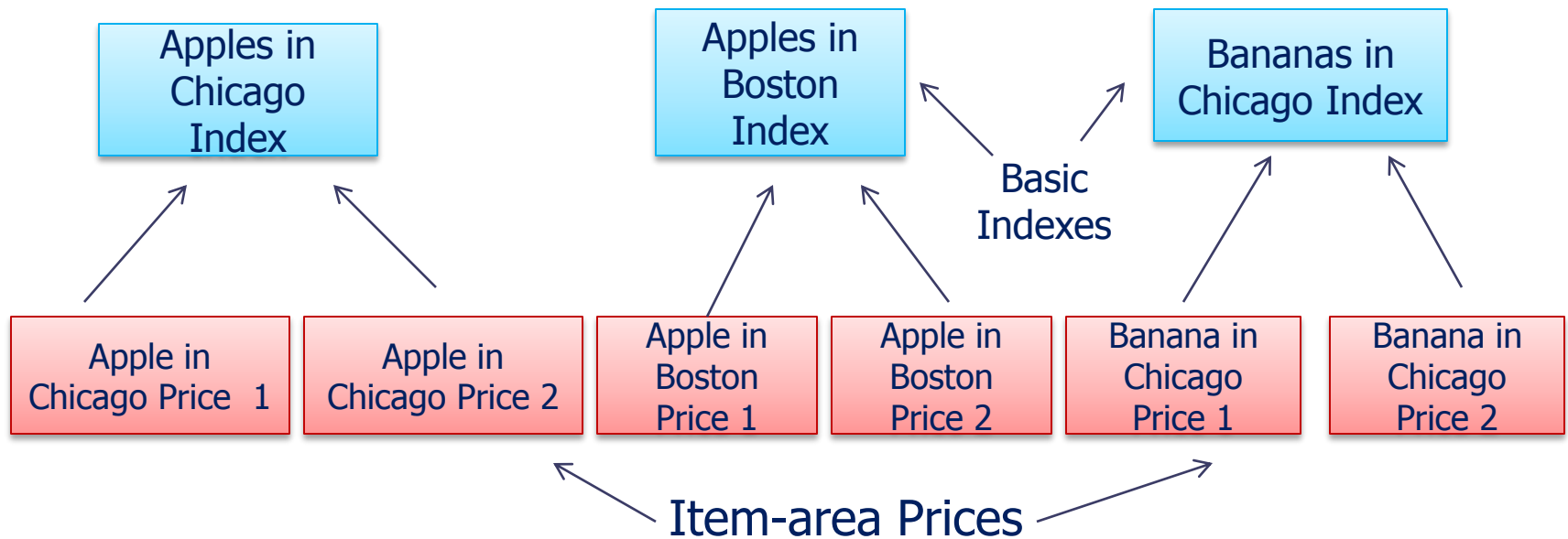
Estimation

- The CPI is created in two stages; the creation of basic indexes and the aggregation of those indexes
- A geometric means formula is used for most items at the first stage; a Laspeyres formula is used in the second stage, except for the C-CPI-U:

Index	Stage	
	1st Stage: Averaging of price changes with each item-area cell	2nd Stage: Aggregation of item area indexes
CPI-U CPI-W	Geometric Mean formula for most items (implicitly measures substitution within item/area cells)	Laspeyres across item/areas (no substitutions)
C-CPI-U	Laspeyres for a few items	'Superlative' Tornqvst index (accounts for observed consumer substitution across items)

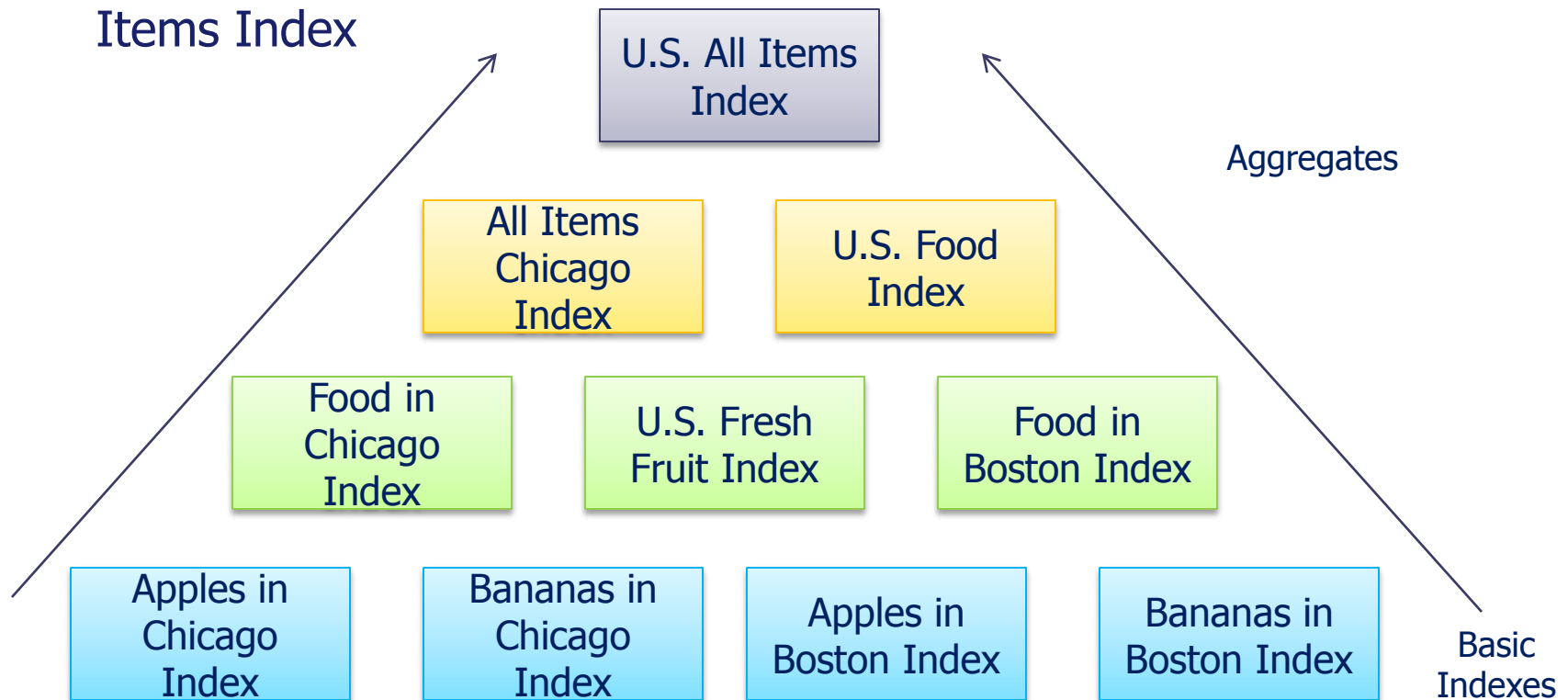
CPI Aggregation Example

- The CPI is calculated in a two-stage process. At the first stage, the price changes within each item-area are averaged to form the item-area or “basic” indexes



CPI Aggregation Example

- At the second stage, these basic item-area indexes can be averaged together to form aggregate indexes for different areas and item categories, including the overall U.S. City Average All Items Index



Estimation

- The Laspeyres formula holds quantity constant, assuming no substitution, and by construction is an upper bound to the change in the cost of maintaining a standard of living.
- The Laspeyres formula is used at the lower level in categories where substitution would be difficult or unlikely, such as medical care services
- The Laspeyres formula is used at the upper level in the aggregation of basic indexes (except in the C-CPI-U); thus there is an assumption of no substitution between fruits and vegetables or gasoline and airline fares

Estimation

Laspeyres formula: (R is the price relative, P_0 and Q_0 are base period prices and quantities, P_t is current period price)

$$R_{t,0} = \frac{\sum P_{t,i} Q_{0,i}}{\sum P_{0,i} Q_{0,i}}$$

Estimation

- The geometric mean formula is used for basic indexes at the lower level for items where some substitution is realistic. It implicitly assumes a degree of substitution among the items used in calculating a price index.
 - ▶ Substitution among different types of steak.
 - ▶ Does not assume substitution of hamburger for steak.
 - ▶ The geometric means formula is not used at all for upper level aggregation in the CPI-U and CPI-W

Estimation

- Geometric means formula: (R is the price relative, P_0 and Q_0 are base period prices and quantities, P_t is current period price, S is the share of the total weight for the item)

$$R_{t,0} = \prod (P_{t,i} / P_{0,i})^{S_{0,i}}$$

- Total product of ratio of current price to base month price raised to the power of item's share

Estimation

- Seasonal adjustment is used to eliminate price change that occurs of about the same time and magnitude every year.
- Seasonal factors are created for many different CPI series and are updated yearly.
- Seasonally adjusted data are typically used for analysis of month-to-month price change. Because it is subject to revision, it should not typically be used for escalation or in contracts.

Treatment of missing price data

- All programs must deal with missing price data
 - ▶ If a price is temporarily or seasonally unavailable, the price is typically imputed and a price will be collected in a future period.
 - ▶ If the item is no longer sold or traded, each price program has procedures for:
 - How to replace the item in the sample
 - What price change to use between the old and new sample items

Treatment of missing prices in the CPI

- If a sampled item is temporarily or seasonally unavailable, and expected to be available in the future, the price is imputed and remains in the sample.
- When a sampled item will no longer be sold in the outlet, CPI field staff “substitute” to the most similar item still for sale in that outlet.

Treatment of missing prices in the CPI

- When the field makes a substitution, the CPI commodity analyst (CA) can:
 - ▶ (1) Deem the substitution “comparable”
 - ▶ (2) Estimate directly the quality change between the two versions (“quality adjustment”)
 - ▶ (3) Deem the substitution non-comparable

Treatment of missing prices in the CPI

- (1) Comparable substitution: If the CA determines there is little or no change between the two versions, she codes the substitution as “comparable.” This means the entire price change between the discontinued and new versions are shown in the index as price change
 - ▶ Implicitly, this also means the quality change between the two versions is zero

Treatment of missing prices in the CPI

- (2) Direct quality adjustment: If the CA determines there is a change in quality between the discontinued and new versions—and the CA can estimate the value of that quality change—the CA will use a quality adjustment for that item
 - ▶ Regression models (“hedonics”) or estimates of production costs may be used to estimate the value of a quality change
 - ▶ For example, if the new car is \$100 more expensive than the older model, but has a stronger bumper that cost the automaker an additional \$10 to install, the CA may only show a \$90 price increase

Treatment of missing prices in the CPI

- (3) Non-comparable substitution: If the CA determines the quality change between the two versions is significant—and cannot estimate the value of that quality change—the CA will code that substitution “non-comparable”
 - ▶ When a quote is coded non-comparable, the price change for that items is effectively imputed by similar items in the same geographic area

Publication

- **Headline CPI indexes**
 - ▶ CPI-U All Items U.S. City Average CPI for urban consumers is the most used and cited CPI.
 - ▶ CPI breakdowns for food, energy, and all items less food and energy often receive attention.
 - ▶ Besides the CPI-U, the BLS publishes the CPI-W, which is based on the spending patterns of the wage-earning subset of the population, and the C-CPI-U.

Publication

- The CPI is broken down by type of item and by geographic location:
 - ▶ Because of small sample size, many geographic-item combinations are not published.
 - ▶ Broader geographic indexes have more indexes published. Also, some individual city CPIs are bimonthly or semi-annual rather than monthly.

Challenges and Future Improvements

- Challenges and improvements in the CPI:
 - ▶ Use of alternative data sources such as corporate and scanner data
 - ▶ Improving Telephone Point-of-Purchase
 - Adjust to new telephone usage patterns
 - Improve response rates and cooperation

Limitations of the CPI

- Sampling error: CPI is based on a series of samples, so the price change of the sample may not be the price change of the entire universe of prices
- Standard error for all items one month change in 2013 was 0.03 percent; for 12 month change it was 0.08 percent
- The standard error for many narrower CPI categories and geographic regions is much higher

Limitations of the CPI

- Consumer substitution: Consumers change spending behavior over time in response to price change and other reasons
 - ▶ Use of geometric means formula assume unitary elasticity of substitution at lower level, probably little or no bias in first stage
 - ▶ CPI-U and CPI-W assume no upper level substitution and still have bias at second stage
 - ▶ Final C-CPI-U should be nearly free of substitution bias

Limitations of the CPI

- Quality change bias
 - ▶ Goal of the CPI is constant quality price change
 - ▶ Many methods of quality adjustment are used but they are imperfect and subject to error
- New goods bias
 - ▶ New goods and new models entering the economy constantly
 - ▶ Biennial weight updates and more frequent sample rotation have improved the timeliness of the CPI market basket, but there is still some time lag

Limitations of the CPI

- Other sources of bias
 - ▶ CPI subject to the biases of survey that feed into it, such as the CE
 - Example: Under reporting of tobacco purchases
 - ▶ Outlet bias: Consumers may substitute to new outlets that offer lower prices
 - ▶ CPI is based on inflation experience of urban consumers rather than entire population

Contact Information

General Web Information:

BLS - www.bls.gov

CPI - www.bls.gov/cpi

Information requests:

(202) 691-7000

CPI_Info@bls.gov



Contact Information

**Steve Reed
Economist**

202-691-5378

reed.steve@bls.gov

