Incorporating a geometric mean formula into the CPI

*Beginning in January 1999, a new geometric mean
formula will replace the current Laspeyres formula
in calculating most basic components
of the Consumer Price Index; the new formula will better
account for the economic substitution behavior of consumers*

Kenneth V. Dalton,
John S. Greenlees,
and
Kenneth J. Stewart

This article describes an important improvement in the calculation of the Consumer Price Index (CPI). The Bureau of Labor Statistics plans to use a new geometric mean formula for calculating most of the basic components of the Consumer Price Index for all Urban Consumers (CPI-U) and the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W). This change will become effective with data for January 1999.[1]

The geometric mean formula will be used in index categories that make up approximately 61 percent of total consumer spending represented by the CPI-U. The remaining index categories, which are shown in exhibit 1, will continue to be calculated as they are currently. On the basis of BLS research, it is expected that the use of the new formula will reduce the annual rate of increase in the CPI by approximately 0.2 percentage point per year.

Background

The Bureau of Labor Statistics has a long tradition of research and operational innovation regarding the measurement of prices. In the past few years, the Bureau has made a series of improvements that reflect the agency's ongoing commitment to keeping its measures on changes in consumer prices as accurate and up to date as possible.

In 1995 and 1996, for example, steps were taken to resolve the problem of "functional form

bias" in the CPI.[2] This upward bias was a technical problem that tied the weight of a CPI sample item to its expected price change. The flaw was effectively eliminated by changes to the CPI sample rotation and substitution procedures and to the functional form used to calculate changes in the cost of shelter for homeowners. In 1997, a new approach to the measurement of consumer prices for hospital services was introduced.[3] Pricing procedures were revised, from pricing individual items (such as a unit of blood or a hospital inpatient day) to pricing the combined sets of goods and services provided on selected patient bills, facilitating the measurement of price change in an area in which the products and services available are undergoing significant change. Effective with the data for January 1998, the Bureau instituted a number of important changes associated with the 1998 CPI revision.[4] Expenditure weights are now based on 1993–95 consumer spending patterns, a new sample of geographic areas was introduced that better represents the current distribution of the U.S. population, and a new classification structure was incorporated that better reflects the categories of goods and services consumers buy. In January 1998, the Bureau also improved its treatment of changes in the quality of personal computers.[5]

The Bureau plans further improvements to the CPI as well. A hedonic model to adjust television prices for changes in quality is slated to begin in January 1999, as is a change in the way utility rebates are treated in the index.[6] The Bureau also

Kenneth V. Dalton is Associate Commissioner for Prices and Living Conditions; John S. Greenlees is Assistant Commissioner for Consumer Prices and Price Indexes; and Kenneth J. Stewart is an economist in the Division of Consumer Prices and Price Indexes, Bureau of Labor Statistics.

| Exhibit 1. | Components of the CPI retaining the arithmetic mean (Laspeyres) formula |

Selected shelter services

Rent of primary residence
Owner's equivalent rent of primary residence
Housing at school, excluding board

Selected utilities and government charges

Electricity
Maintenance of residential water and sewerage
Telephone services, local charges
Utility natural gas service
State and local registration, license, and motor
  vehicle property tax
Cable television

Selected medical care services

Physicians' services
Eyeglasses and eye care
Hospital services
Dental services
Services by other medical professionals
Nursing homes and adult day care

plans to improve the timeliness of future CPI revisions. In this regard, the size of the Consumer Expenditure (CEX) Survey, which is used as the source of the market basket weights for the CPI, will expand by approximately 50 percent, and the lag between the collection of information on expenditures and the introduction of the market baskets into the CPI will decrease.[7] Further, in addition to supporting more timely updates of CPI expenditure weights, the expanded sample for the CEX Survey will support the publication of production-quality superlative indexes. Under certain conditions, superlative indexes account for substitution across CPI item categories in response to relative price changes. Finally, the Bureau plans to improve its procedures for selecting outlet and item samples, to allow for more frequent sample updates for those consumer markets undergoing rapid change.[8]

The geometric mean formula

The CPI currently is constructed using a set of constant, or fixed, implicit quantity weights. Thus, the index does not reflect the fact that consumers can—and do, to some degree—insulate themselves from the impact of higher prices by adjusting their spending to favor relatively lower priced goods or services. Consequently, compared with a measure that reflects this substitution effect, the current CPI tends to overstate the rate of price increase consumers experience.

In contrast to the fixed quantity weights of the current CPI arithmetic mean formula, the geometric mean estimator employs a set of fixed expenditure proportions as weights to be used in averaging the prices of individual items within a CPI basic index. Fixing the relative expenditure proportions, rather than the relative quantities, implies that consumers can alter the quantities of goods and services they buy—albeit within the narrow range of a CPI category—when the relative prices of those goods and services change.

The CPI is constructed as an aggregation of basic indexes computed for approximately 200 categories of items, such as "ice cream and related products," in each of 38 geographic areas. Within each of these index components, or strata, prices for specific items in a sample of outlets (stores) are combined to produce a basic index. As noted above, the geometric mean formula will be used solely to average prices within the item-area strata. Consequently, the use of the formula will address only the issue of consumer substitution within strata.[9]

The research the Bureau has conducted over the past few years strongly suggests that the use of the geometric mean estimator at the basic level of index construction in the CPI will produce a measure that more accurately reflects the impact that changing prices have on the average U.S. household. The remainder of this article presents (1) a summary of the factors that led the Bureau to consider using the new formula, (2) the empirical research the Bureau conducted in evaluating the formula, and (3) the specific BLS decisions that were reached in regard to using the formula in the CPI.

The possibility of using the geometric mean formula to calculate basic indexes in the CPI was first raised by BLS researchers in 1993.[10] One initial motivation for research into this area was the problem of *functional form bias* (sometimes referred to as *formula bias* or *elementary index bias*) then present in the CPI. Functional form bias occurred because of technical problems involved in using observed expenditure information to estimate the quantity weights employed in the index formula. Because the geometric mean formula does not require quantity data, it does not suffer from this problem. As noted above, functional form bias was eliminated from the CPI in a series of steps, beginning in January 1995 and ending in July 1996, using an approach that did not involve the geometric mean formula. The Bureau, however, continued research on the formula and developed an experimental geometric mean index, in part due to a recognition that the formula offered a means of reflecting consumer substitution behavior.

In December 1996, the Advisory Commission to Study the Consumer Price Index, commonly known as the Boskin Commission, recommended the use of the geometric mean formula for the aggregation of prices within all categories of items in the CPI.[11] This recommendation was based upon the belief that a geometric mean formula would help to correct what the Commission called "substitution bias."

Substitution can take several forms, corresponding to the types of item- and outlet-specific prices used to construct the basic indexes:

- Substitution among brands of products, for example, between brands of ice cream
- Substitution among sizes of products, for example, between pint and quart packages of ice cream
- Substitution among outlets, for example, between a brand of ice cream sold at two different stores
- Substitution across time, for example, between purchasing ice cream during the first or the second week of the month
- Substitution among types of items within the category, for example, between ice cream and frozen yogurt
- Substitution among specific items in different index categories, for example, between ice cream and cupcakes

Thus, in response to an increase in the price charged by a store for a certain brand of ice cream, a consumer could respond by redistributing his or her purchases along any of several dimensions represented by other priced items in the category: to another brand of ice cream whose price had not risen, to a larger package of ice cream with a smaller price per ounce, to ice cream at a different store where ice cream is on sale, or to a brand of frozen yogurt. The consumer also could respond by postponing the ice cream purchase until a later date. (Prices for CPI items are collected throughout the month and then averaged.)

Finally, the consumer could substitute a specific alternative dessert item, such as cupcakes or apples, that is in another CPI category. This form of substitution, although across CPI categories, would still have the effect of reducing the quantity consumed of the higher priced brand of ice cream relative to the quantities consumed of other items within the ice cream stratum. Like the other forms of substitution, such cross-category substitution is implicitly addressed by the use of the geometric mean formula. Note, however, that the formula does *not* address *overall* substitution across categories, such as between ice cream products in general and apples in general. Thus, the geometric mean formula will not be used to combine the basic indexes in the CPI, such as those for ice cream products and apples, into the overall index. In the same way, the use of the geometric mean formula within categories does not address the issue of whether consumers can, or do, respond to a general increase in the price of ice cream products by, for example, forgoing dessert altogether.

Note also that the formula only allows for a *degree* of substitution by consumers as a group: while the formula implicitly assumes that at least some customers will change their purchasing patterns when the relative prices of items change, it is not inconsistent with some individual consumers continuing to purchase their favorite products even when they become relatively more expensive.

In April 1997, the Bureau began issuing an experimental CPI that used the geometric mean estimator in the calculation of all basic components of the index. This experimental index, called the CPI-U-XG, afforded a quantitative estimate of the impact that the use of a geometric mean formula at the lowest level of index construction would have on the performance of the CPI. At that time, the Bureau stated that it would continue research on the question of using a geometric mean estimator in the CPI.[12] This research would entail a thorough review of the theoretical literature concerning different index number formulas, a review of the practices of other countries and the justifications for their choices of formula, and a systematic and comprehensive analysis of available data to evaluate the appropriateness of using the geometric mean formula in the construction of each basic component of the CPI.

One of the principal objectives of the BLS research on geometric means was to determine, for each of the basic index categories that make up the CPI, the extent to which consumers can be expected to alter their spending when relative prices change. Three types of evidence were examined in this effort: (1) highly detailed supermarket scanner data on prices charged for, and quantities sold of, items in a limited number of categories; (2) measures of the extent of substitution at index calculation levels above the basic level, which can be viewed as providing indirect evidence concerning the likelihood of substitution behavior within item categories; and (3) estimates of the magnitude and prevalence of the substitution effect derived from a survey of the relevant empirical literature. For a variety of reasons, none of the three types of evidence provided definitive support regarding the existence and magnitude of the substitution effect in each of the basic index categories. For example, although analysis of the detailed data on prices and quantities strongly indicated the existence of a significant substitution effect, the analysis was possible for only a few categories of items because of limited availability of the data.

Taken in its entirety, however, the evidence unambiguously supported the proposition that consumers can, and do, alter their purchasing behavior in response to changes in the array of prices that they confront in the marketplace. As stated earlier, the geometric mean estimator can better reflect the effects of such changes in consumer spending than can the current CPI formula.

Incorporating the formula into the CPI

On the basis of the preceding analysis of the evidence, the use of the geometric mean estimator at the basic level of index calculation in the CPI can be expected to produce an overall index that better reflects the impact that changing prices have on the average consumer. Consequently, the Bureau plans to use the geometric mean estimator in most CPI basic indexes, beginning with data for January 1999. The basic indexes that will continue to be calculated as they are now represent important exceptions to this general rule. The excepted indexes

are shown in exhibit 1 and fall into three general categories, the rationales for whose exclusion are presented next.

Selected shelter services. The total residential housing stock changes slowly. Thus, consumers as a group cannot freely alter their purchases of housing services in response to changes in the relative prices of different rental units. Consequently, the item categories of residential rent, owner's equivalent rent, and housing at school will not employ the geometric mean estimator. As of December 1997, these three categories together had a relative importance of 27.315 percent in the CPI-U and 25.500 percent in the CPI-W.

Selected utilities and government charges. The services in these categories are provided primarily by governments or by regulated monopolies. Although CPI pricing samples within a basic index may include different service providers, in most cases consumers can substitute between these providers only by moving into a different service area. The BLS analysis of the substitution effect above the basic index level supports the continued use of the current formula in the utilities sector. Additionally, estimated utility demand elasticities obtained from the economics literature mentioned before are low, at least in the short run, and thus argue for maintaining the current approach. The six categories under selected utilities and government charges had a total relative importance of 7.067 percent in the CPI-U and 7.347 percent in the CPI-W, as of December 1997. (These relative importance totals include the prorated weight of two unpublished indexes, unsampled motor vehicle fees and unsampled video and audio, whose movements will be imputed by a mixture of geometric mean and arithmetic mean indexes.)

Selected medical care services. The decision to retain the current formula for professional and hospital medical care services is based largely on the results of the survey of demand elasticities in the economics literature. The relatively large elasticities for some prescription drugs contrast with low elasticities for hospital and professional services. Insurance firms and health plans, acting as agents for consumers, may be more sensitive to relative price changes than are consumers themselves, who pay only a share of the cost directly. It is therefore difficult to justify the proposition that consumers are highly responsive to relative prices in their direct purchases of health care services. The six excluded medical care categories (along with corresponding unpublished health insurance indexes) had a total relative importance of 4.159 percent in the CPI-U and 3.484 percent in the CPI-W, as of December 1997.

These exclusions leave three major groups of the CPI—food and beverages, apparel, and other goods and services—that will use the geometric mean formula exclusively in the calculation of basic indexes. The other major groups each will have one or more excepted categories. Within the transportation group, State and local registration, license, and motor vehicle property tax charges will continue to use the current formula. Within the recreation group, the cable television index will continue to be calculated as it now is, as will the telephone services local charges index within the education and communication group. The housing group will have six important exceptions to the use of the geometric mean estimator, three within the shelter category and three within fuels and utilities. In the medical care group, the medical care commodities indexes will employ the geometric mean formula, while most indexes for medical care services will continue to use the current formula.

It is important to note that, in addition to the evidence described above, the Bureau considered other factors in reaching its decisions on the use of the geometric mean estimator. Two of these additional considerations are worthy of mention. First, the geometric mean estimator does not suffer from the persistent functional form bias that affected the current formula before the special, corrective procedures were implemented in 1995 and 1996. Use of the geometric mean estimator as outlined above obviates the need to employ these special procedures in all but the several basic indexes that will continue to use the current formula. Largely eliminating the need for special procedures will, for the most part, make unnecessary the continued monitoring of how well they are achieving their intended objective. Second, index calculation formulas often are evaluated with respect to how well they satisfy certain performance criteria, often referred to as tests. Within the context of the list of tests generally imposed on price indexes, the geometric mean formula performs at least as well as the current, Laspeyres index.

The geometric mean estimator will be introduced into both the CPI-U and the CPI-W effective with data for January 1999, in accord with the past practice of introducing methodological changes at the beginning of a calendar year. The Bureau will continue to publish "overlap" CPI-U and CPI-W series using the current calculation method for the first 6 months of 1999. These indexes will not be published regularly for months subsequent to June 1999, but will be available upon request.

The experimental CPI-U-XG, using the geometric mean estimator in the calculation of all basic indexes and issued since April 1997, will be calculated through data for December 1998. Users of this series are reminded that the implementation of the geometric mean formula in 1999 will not apply to all basic indexes. Therefore, the movements of the CPI-U-XG relative to the CPI-U during 1998 may not be indicative of the future impact of the geometric mean formula in the CPI. □

Footnotes

[1] See, for example, "Improving CPI sample rotation procedures," *Consumer Price Index Detailed Report* (Bureau of Labor Statistics, October 1994), pp. 7–8.

[2] Elaine M. Cardenas, "Revision of the CPI hospital services component," *Monthly Labor Review*, December 1996, pp. 40–48.

[3] John S. Greenlees and Charles C. Mason, "Overview of the 1998 revision of the Consumer Price Index," *Monthly Labor Review*, December 1996, pp. 3–9.

[4] "Using a hedonic model in the CPI to adjust personal computer prices for changes in quality," *Consumer Price Index Detailed Report* (Bureau of Labor Statistics, June 1997), p. 18.

[5] See "Using a hedonic model to adjust television prices in the Consumer Price Index for changes in quality," *Consumer Price Index Detailed Report* (Bureau of Labor Statistics, June 1998), p. 5; and "Improvements to CPI procedures for handling utility refunds," *Consumer Price Index Detailed Report* (Bureau of Labor Statistics, July 1998, p. 5.

[6] Testimony of Katharine G. Abraham, Commissioner of Labor Statistics, before the Subcommittee on Human Resources of the House Committee on Government Reform and Oversight, Apr. 29, 1998.

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

[8] Recent empirical analyses of substitution across strata are presented in, for example, Ana M. Aizcorbe and Patrick C. Jackman, "The commodity substitution effect in CPI data, 1982–91," *Monthly Labor Review*, December 1993, pp. 25–33; and Matthew D. Shapiro and David W. Wilcox, "Mismeasurement in the Consumer Price Index: An Evaluation," in Ben S. Bernanke and Julio J. Rotemberg, eds., NBER *Macroeconomics Annual, 1996* (Cambridge, MA, MIT Press, 1996), pp. 93–142.

[9] Brent R. Moulton, "Basic components of the CPI: estimation of price changes," *Monthly Labor Review*, December 1993, pp. 13–24. Among other recent papers that propound theoretical or empirical results concerning the geometric mean formula are W. Erwin Diewert, *Axiomatic and Economic Approaches to Elementary Indexes*, Working Paper No. 5104 (Cambridge, MA, National Bureau of Economic Research, May 1995); Marshall B. Reinsdorf and Brent R. Moulton, "The Construction of Basic Components of Cost-of-Living Indexes," in Timothy F. Bresnahan and Robert J. Gordon, eds., *The Economics of New Goods* (Chicago, University of Chicago Press, 1997); and Brent R. Moulton and Karin E. Smedley, "A Comparison of Estimators for Elementary Aggregates of the CPI." Paper presented at the Western Economic Association International Conference, San Diego, July 7, 1995.

[10] Advisory Commission to Study the Consumer Price Index, *Toward a More Accurate Measure of the Cost of Living*, Final Report to the Senate Finance Committee, Dec. 4, 1996.

[11] *The Experimental CPI Using Geometric Means (CPI-U-XG)* (Bureau of Labor Statistics, Apr. 10, 1997).

Bureau of Labor Statistics Internet

The Bureau of Labor Statistics World Wide Web site on the Internet contains a range of data on consumer and producer prices, employment and unemployment, occupational compensation, employee benefits, workplace injuries and illnesses, and productivity. The homepage, at **http://stats.bls.gov** can be accessed using any Web browser.

Also, some data can be accessed through anonymous FTP or Gopher at **stats.bls.gov**