THE CONSUMER PRICE INDEX

Philip J. McCarthy *New York State School of Industrial and Labor Relations*
Cornell University

ALMOST EVERYONE worries about the prices he pays for all kinds of things. Perhaps the best way for an individual to find out how the price level changes is to follow in newspapers the Consumer Price Index (CPI) of the Bureau of Labor Statistics, a part of the U.S. Department of Labor. The individual consumer sees the CPI as a good measure of price changes in goods and services that he purchases. He reacts to newspaper statements such as "In September of 1969, the average urban family must spend $12.93 for the same amount of goods and services that could be obtained for $10.00 in 1957–59." He wonders whether or not his income has increased sufficiently to compensate for this increase in prices.

Unions and management pay particular attention to the CPI because they know its value will play a critical role in wage agreements and that a 7% annual increase, for example, in the CPI will lead to a demand for at least

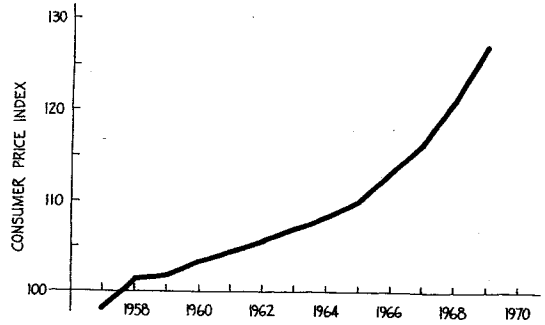

FIGURE 1
Consumer Price Index, annual averages (1957–59 = 100)

a 7% increase in wages. Furthermore, in January of 1969 there were 2.55 million workers whose wages were covered by contracts containing *escalator* provisions, that is, provisions calling for automatic changes in wage rates in accordance with specified changes in the CPI.

Finally, economists concerned with the fiscal and monetary policies of the U.S. Government use the CPI as one of the principal indicators of the existence of an inflationary spiral in which higher prices lead to higher wages, which, in turn, lead to a greater demand for goods, thus raising prices again, and so on. During recent years, most writings on the state of the U.S. economy have emphasized the fear of inflation as evidenced by the increasing values of the CPI. The reason for this fear is illustrated clearly in Figure 1, which shows annual average values of the CPI for 1957–69. These values have been rising at an increasing rate since about 1964.

Although statistical studies of prices and living conditions in the U.S. were conducted in the late nineteenth and early twentieth centuries, the first complete "cost-of-living" indexes were published by the Bureau of Labor Statistics in 1919. They referred to 32 large shipbuilding and industrial centers, and arose through an agreement between the Shipbuilding Labor Adjustment Board and labor chiefs that one of the factors to be considered in settling labor disputes was that of "adjusting wages to the higher cost of living resulting from the war."

Since that time, the CPI has broadened in scope and increased in importance. This has led to many professional appraisals of the CPI, among them one by an advisory committee of the American Statistical Association in 1933–34 and one by the National Bureau of Economic Research in 1959–60. These appraisals have influenced major revisions in the CPI, made at approximately ten-year intervals since 1940, as well as many minor revisions.

PROBLEMS IN CONSTRUCTING THE CPI

The typical consumer is quite aware of changes that occur in the prices of goods and services that he purchases regularly. He knows when the price of gasoline increases by one cent per gallon, or when the price of milk is increased by one cent per quart. Furthermore, he is able to predict with reasonable accuracy the impact of these price changes on his monthly or yearly budget, and he is therefore in a position to estimate the extra income (including an allowance for taxes) he must obtain in order to compensate for these increases. Price changes relating to less frequent and more sporadic purchases—clothing, doctors' services, appliances, automobiles, and homes, for example—are not as visible to the individual consumer, and it is much more difficult for him to assess their effects on his budget. Thus he may well know that it costs him more to live, or less to live, during the current year than in the preceding year, but he would find it almost impossible to provide an exact value for this change in his cost of living. His problem would be further complicated by changes that might occur in such matters as family status (births and deaths), family desires (color TV instead of black and white), and family purchasing patterns (chicken instead of steak). And yet it is exactly this evaluation—for the nation as a whole, for separate regions and cities of the nation, for different classes of expenditures (food, clothing, housing, transportation, and the like), and on a monthly basis—that the CPI provides. This essay is concerned primarily with the contribution of statistics to the construction of the CPI.

If the individual consumer actually wished to keep sufficiently detailed records to measure the change in prices of the items he purchases from, say, January of 1968 to January of 1969, he might proceed as follows. During January of 1968, he would keep a record of every purchase and would then summarize these purchases in terms of quantities (numbers of quarts of milk, number of pairs of shoes, number of haircuts, and the like) and the unit price of each item. In effect, he defines a "market basket" of goods and services and obtains the cost of this market basket in January 1968. Suppose this cost is \$750.00. In January 1969, he would price exactly the same market basket of goods and services. Let us assume that the cost then was \$800.00. Therefore, this consumer's personal CPI for January 1969 with January 1968 weights (quantities) and with January 1968 as base ($= 100$) is $100 \times 800/750 = 107$. The CPI is effectively computed in the same way, but the operation is much more complicated, since it deals with a large population rather than a single consumer.

Even if we assume that the family status and desires of our consumer did not change from January 1968 to January 1969—an assumption that can seldom be exactly true—there are many practical problems and tantalizing

questions that would plague those responsible for the construction and interpretation of this index. In particular,

(1) Even for a single consuming family, the number of different items purchased during a month or year is extremely large, and the continual pricing of such a list would be costly and time-consuming. Preparation of a CPI for groups of families obviously magnifies this problem to a wholly impractical size.

(2) The average family has available to it a host of vendors of goods and services, and prices vary from vendor to vendor. In January of 1969 should one attempt to return to the same places where purchases were made in 1968? And what about stores that have gone out of business or stores that have come into existence in the interim?

(3) Prices vary from day to day. Must our consuming family return to the same store at the same time to compute the cost of its market basket in January of 1969? And how does one take into account the facts that some families take advantage of sales and others do not and that some families shop on a day-to-day basis while others shop less frequently?

(4) Suppose that a major purchase, such as an automobile, was made by our family in January 1968. The fixed market basket still contained this item in January 1969 even though such a purchase was not made at that time. How should this be handled?

(5) The goods and services available to the consumer change from time to time. Items contained in the original market basket may not exist in the stores when we return to price them at a later date, or they may exist only at an improved or lowered quality level. What do price changes mean under these circumstances?

(6) We also observe that our consuming family—even though its composition, status, and desires are assumed not to have changed—may be able to "beat" a rise in prices by appropriately altering its purchasing patterns. Suppose, for example, that they notice that the price of steak has risen sharply and decide that the desires of the family are as well satisfied by chicken as by steak. If chicken is substituted for steak, the family's satisfaction level will not change, and yet the amount spent for food may actually decrease in spite of what may be a general increase in the price level. In effect, we can think of replacing the "fixed-contents" market basket with a "fixed-satisfaction" market basket.

The above problems relate to both measurement and concept. Thus we may find that a particular item of merchandise has changed in quality, and yet it may be a most difficult task to measure this change in quality and to translate it into a dollar value. The Bureau of Labor Statistics does attempt to account for these changes in quality in constructing the CPI.

Item 6 above is at a deeper conceptual level. Most economists would prefer a "constant-utility" or "constant-satisfaction" or "welfare" type of price index to the "fixed-market-basket" type of index that is currently produced by the Bureau of Labor Statistics. In other words, what change in expenditures must a consuming family make from one time to another in order to maintain a constant level of satisfaction, with the recognition that the contents of the market basket can be changed in order to accommodate changes in prices, changes in products and the like? Although steps are being taken to move the CPI in the direction of such a "constant-utility" index, the practical and conceptual problems are difficult to overcome, and progress has been slow. As a matter of fact, in September 1945, the official title of the index was changed from "The Cost of Living Index" to "The Consumer Price Index" in order to emphasize the distinction between these two approaches. In this essay we shall not treat these conceptual problems, but rather we shall focus on the role of statistics as it helps to solve CPI measurement problems.

SAMPLING AND THE CPI

Many of the above problems become more manageable if attention is shifted from the individual consuming family to a group, or *population,* of consuming families. Thus, even though some members of the population did purchase automobiles in January 1968, other members of the population did not make such a purchase, and similarly for January 1969. An automobile can then be introduced into the market basket with a weight that will reflect the effects of changes in its average price, averaged over the population of families. This shift in emphasis from the individual consumer to a population of consumers, where differences among the members of this population are known to exist, means that the construction of the CPI requires statistical sampling (and analysis) from the population of consuming families.

Moving in this direction also forces one to think in terms of the *population of goods and services* available to all members of the population of consumers and in terms of the *population of outlets* at which all of these goods and services may be purchased. Furthermore, it is manifestly impossible to study every consuming family, and to price each item of consumption in all the outlets where it can be purchased. Hence it becomes necessary to select samples from each of these populations and to draw inferences from the samples to the entire populations. The methods of statistics can assist and guide these steps.

The Consumer Price Index has never attempted to measure the changes in the prices of goods and services for all families and individuals living in the United States. Rather, because of its traditional use in collective bargaining between labor and management, its scope has been restricted to urban

families. More specifically, the *population of consuming families* covered by the index consists of all urban families (including single workers) for which 50% or more of the family's income comes from wages or from salaries earned in clerical occupations and for which at least one member of the family unit works for at least 37 weeks during a year.

In determining the contents of an average, or "representative," market basket of goods and services for this defined CPI population of consuming families, it is impossible to study every family. Hence a sample, or a portion of this population, must be chosen. Because a list of the members of the population is not available, complex methods of statistical sampling must be employed.

In brief, a sample of urban communities is selected first, and then a sample of families is taken from each of the selected communities. In the 1959–64 revision, this was accomplished through an original selection of 50 cities. The 12 largest cities in the U.S., plus one city from Alaska and one city from Hawaii, were automatically included in this sample of cities. The remaining urban areas of 2500 and over were placed in homogeneous groups according to size and geographic location and 36 cities were chosen from these groups in accordance with probability methods of selection. An additional six large cities were added to the sample later, primarily so that individual city indexes could be published. This sample of 56 cities serves not only as a basis for studying the expenditure patterns of consuming families, but also as a basis from which to obtain the prices that must be used to determine the current cost of the CPI's market basket of goods and services. The CPI will continue to be based upon this sample of 56 cities until the next major revision takes place.

Within each of the 56 chosen cities, a sample of consuming units was selected and interviewed during the period 1960–61. These units were drawn in accordance with the tenets of statistical sampling theory; the goal was to choose samples in such a way that objective measures could be devised for assessing the likely size of deviations between averages for the sample and corresponding (unknown) averages for the whole population. This particular survey was called the *Consumer Expenditure Survey* (CES) because its primary purpose was to collect data relating to family expenditures for goods and services used in day-to-day living. Information was obtained through lengthy personal interviews with family members. Among the items recorded during this interview were the following:

(1) A complete account of receipts and disbursements for the preceding calendar year.
(2) The estimated value of goods and services received free.
(3) The characteristics of housing occupied by both home owners and renters.

(4) An inventory of major household furnishings owned.
(5) A detailed listing for a seven-day period of expenditures for food and beverages, household supplies, and tobacco.

Altogether, intensive interviews were conducted with 9476 consuming units. Of these, 4860 interviews were with members of the defined population of consumers for the CPI. The remaining 4616 units did not satisfy the CPI restriction to the population of consuming families, although the data obtained from them are of value for other purposes.

Data collected in the CES interviews with members of the CPI population were used for a variety of purposes in determining the structure of the current CPI. First of all, these data determined the complete contents and total cost of each family's market basket in the survey year. The number of items of expenditure in all market baskets totaled about 1800. The items were classified into five major groups: food, housing, apparel and upkeep, transportation, and health and recreation. Each of these groups was further subdivided, so that the final classification scheme consisted of 52 expenditure classes. Some examples of expenditure classes are: meats, eggs, fuel and utilities, housekeeping services, footwear, auto repairs and maintenance, hospital services and health insurance, and reading and education. In effect, the consumer's market basket was divided into 52 compartments. The total cost of each compartment was then determined for all sample consumers selected from within a particular city, and these total costs were expressed in relative terms; for example, at December 1963 prices, meats were estimated to account for 4.45% of the total cost of the nationwide market basket. Different market baskets are used in different cities to allow for differences in such characteristics as climate and the availability of different foods. Thus we use many average market baskets rather than a market basket that applies to a single consuming unit.

As observed earlier, the CES interviews provided a market basket filled with some 1800 items that consumers had purchased. It would be impossible to price all of these items in their almost infinite variety each month, even recognizing that this pricing need be carried out only in the 56 index cities. Hence another sampling problem arises. Again the theory of statistical sampling was used to select a sample of items from each of the 52 expenditure classes. The final sample for the 1959–64 revision contained 309 items. The sampling approach allows the contents of the market basket to change. For example, if one item disappears from the market, replacements may be made by further sampling. The compartments and their original weights, however, remain fixed through time.

Once a sample of items has been selected and its members specified in detail, another, more difficult sampling problem must be faced. Within each of the 56 index cities, prices of the sample items must be obtained on a

monthly basis from the outlets patronized by members of the CPI population of consuming units. It is impractical to obtain price quotations from all possible outlets, and so a sample of outlets must be selected to serve as sources of price information.

Among the problems that have to be considered in developing a sample of outlets are the following:

(1) Ideally, we would define a separate population of outlets for *each* item to be priced and would select a sample from *each* of these outlet populations. This approach would be too costly because even a sample of only four or five outlets for each item would require that 1500 outlets in a city would have to be priced. Furthermore, it is difficult to compile lists of outlets on an item-by-item basis because of merchandising patterns. For example, department stores sell a tremendous variety of items, including clothing, appliances, and furniture; tire and automobile accessory stores also well appliances and toys. Hence compromises must be made in developing a sample of outlets.
(2) The sampling problems are quite different for different items. We can obtain the price rate for electricity by merely visiting the local utility company or the price of newspapers by calling the local publishers. On the other hand, the price of items such as meats and fresh produce vary widely from one grocery store to another, and a fairly large sample of price quotations are required in order to determine the average price of a grocery item with any degree of precision.
(3) There is no clearcut way of identifying the particular outlets that are patronized by the population of consuming families to which the CPI is supposed to apply. The Bureau of Labor Statistics recognizes that the shopping patterns of consuming units now range over a wide geographic area, even though the residences may be confined to a city. Thus the final sample attempts to give proper representation to the downtown and neighborhood areas of the central cities, as well as to suburban areas where so many shopping centers have been developed in recent years.

The field operations of monthly pricing are intricate. Not only must a large field staff be trained and supervised, but one must also obtain the cooperation of store managers, and make provision for businesses that cease operations or come into existence during the ten year period that ordinarily elapses between major revisions. The magnitude of these endeavors is indicated by the fact that food prices are collected in 1775 stores each month and that the Bureau is in touch with about 16,000 outlets for the pricing of nonfood commodities and services.

There is a final sampling problem, not always recognized as such, associated with the CPI. Although it is published monthly, it does not refer to any

definite date during the month. The pricing operation has to be almost continual, and it is therefore necessary to choose a sample of points in time at which the prices are obtained. This sampling is not carried out in as formal a manner as are the other sampling operations. Nevertheless, every attempt is made to ensure, for example, that sale and nonsale days for food are represented in their proper proportions, and that a similar balance is maintained for other items, such as newspapers and theater admissions, whose prices may change periodically.

One important goal in statistical design and analysis is to have an objective measure of the precision of sample analyses. Although this goal has not been fully realized in the complex setting of the CPI, substantial progress was made in this direction during the 1959–64 revision, and crude estimates of sampling error have been obtained. Since 1967, these estimates are given in *The Consumer Price Index,* a monthly bulletin published by the Bureau of Labor Statistics. A recent issue of this bulletin states that ". . . any particular (month-to-month) change (in the CPI) of 0.1 percent may or may not be significant. On the other hand, a published change of 0.2 percent is almost always significant, regardless of the time period to which it relates." All indications are that the sampling operations are reasonably well under control, and that uncertainty in the value of the index due to sampling is relatively small compared to the uncertainties arising from other aspects of the process, for example, the effects of quality changes on the index.

SUMMARY

The production of monthly values of a Consumer Price Index by the U.S. Bureau of Labor Satistics is a highly complex undertaking that involves problems of *basic economic theory* (e.g., choice between a price index or a constant-satisfaction index), *measurement and quantification* (e.g., of changes in the quality of items purchased by consumers), *sampling statistics* (definition of, and selection of samples from, a wide variety of populations), and *operations* (e.g., training and supervision of price reporters).

The Index is concerned with a *population of consuming families* and with the *population of cities* in which these families live. A sample of cities serves two purposes. First, from within the selected cities a sample of consumers can be chosen from which it is possible to determine average expenditure patterns. Second, prices are collected in the selected cities to determine the value of the current CPI. The Consumer Expenditure Surveys also define a *population of goods and services* for which consuming families spend their income. From this population a sample must be taken for current pricing purposes. Within each of the index cities there exist *populations of outlets* at which items can be purchased and samples must be chosen to represent these populations. Finally, there is a *population of times* within a month at which price quotations can be obtained, and this population also must be sampled.

It is certainly true that no one is completely satisfied with the CPI in its present form. Improvements can and probably will be made in many parts of the Index, for example, in the basic data on consumer expenditure patterns, in the sampling of outlets and in the collection of price data from these outlets, in the preparation of indexes for a wider variety of subpopulations, and in techniques for handling quality change problems. There will be continuing pressure also to move the CPI more in the direction of a constant-utility type of index. It may even happen that some completely new approach to the construction of the index may be developed, possibly through the use of newer mathematical techniques. No matter what its form, however, the CPI undoubtedly will remain one of the main indicators of the state of the U.S. economy.

REFERENCES

Ethel D. Hoover. 1968. "Index Numbers: II. Practical Applications." David Sills, ed., *International Encyclopedia of the Social Sciences,* vol. 7, pp. 159–165. New York: Macmillan and Free Press.

Philip J. McCarthy. 1968. "Index Numbers: III. Sampling." David Sills, ed., *International Encyclopedia of the Social Sciences,* vol. 7, pp. 165–169. New York: Macmillan and Free Press.

Erik Ruist. 1968. "Index Numbers: I. Theoretical Aspects." David Sills, ed., *International Encyclopedia of the Social Sciences,* vol. 7, pp. 154–159. New York: Macmillan and Free Press.

U.S. Department of Labor. *The Consumer Price Index: History and Techniques.* Bulletin No. 1517. Washington: U.S. Government Printing Office.