

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Price Statistics of the Federal Government

Volume Author/Editor: Price Statistics Review Committee

Volume Publisher: NBER

Volume ISBN: 0-87014-072-8

Volume URL: <http://www.nber.org/books/repo61-1>

Publication Date: 1961

Chapter Title: Problems Common to the Indexes

Chapter Author: Price Statistics Review Committee

Chapter URL: <http://www.nber.org/chapters/c6483>

Chapter pages in book: (p. 31 - 50)

III

PROBLEMS COMMON TO THE INDEXES

Certain problems are encountered in most or all of the price indexes we are reviewing. We treat these problems generally rather than discuss each problem in connection with each index, indicating special adaptations necessary to particular indexes.

1. FREQUENCY OF REVISION OF THE WEIGHT BASES

The practice of the price collecting agencies in revising the weight bases for the price indexes varies widely. The export and import prices are calculated as the geometric means of the indexes based upon weights of the given and preceding year. The construction cost index is implicitly a given year weight index. These are the only indexes not employing a fixed initial year weight base (Laspeyres index). The weight base of the Wholesale Price Index is generally revised on the basis of the most recent Census of Manufactures—the current weight base is 1954; the next weight base will be 1958. The Consumer Price Index is based upon 1952 weights—the previous base (aside from interim adjustments in 1951) was 1934-36 and the next base will be 1960-62. The Index of Prices Paid by Farmers has a 1955 weight base, that of Prices Received by Farmers a 1953-57 base. In these last two cases, the indexes were recomputed back to 1952.

A fixed-weight base is practically unavoidable so long as elaborate expenditure surveys or production censuses are necessary to provide the weights. The principle on which one should decide when to revise the weight base is that a revision is necessary when the weight base has changed appreciably. In a stable society, revisions could be extremely infrequent; in the rapidly changing American economy, a revision once in a decade or more (as has more than once been the case with the Farm Indexes and the Consumer Price Index) is too infrequent. The rapid pace of introduction of new products in the United States, the large demographic changes in recent decades, the revolution in production methods—these are instances of the changes that dictate frequent revision of weight bases.

If budgetary limitations prohibit frequent weight revisions, the practice of the Agricultural Marketing Service in revising the index back over part of the period since the previous revisions seems preferable to the BLS practice of treating the previous indexes as immutable. Such backward revisions will be improved if each year's weights are an average of initial and terminal year weights (combined with relative weights proportional to the propinquity of the given year to each terminal year).

Quite aside from the fact that we believe the United States can afford good price indexes, however, the budgetary restrictions are being relaxed by the improvements in survey methods for obtaining the basic information. It has become feasible to measure more fre-

quently the changes in weights of at least the larger categories (such as electrical appliances in the CPI) and in the most rapidly changing detailed classes of weights.

We therefore believe that there should be an established program of periodic, comprehensive revision of the weights of the Consumer Price Index and the Farm Indexes at least once every decade (this is an outer limit on the period between complete revisions). In addition there should be smaller surveys (and weight revisions) at least every 5 years concentrating upon the more volatile categories of weights.

Even frequent revisions of weights, however, will not eliminate the need for frequent current substitutions of goods because of the continual disappearance of old and appearance of new goods, for reasons to which we now turn. The "fixed market basket" approach does not demand, or even allow, a strict identity in all goods between weight revisions.

2. SPECIFICATION VS. UNIT VALUE PRICING

In 1934 the Bureau of Labor Statistics adopted "specification" pricing, and since then has sought to price narrowly defined commodities and services to obtain price relatives for price indexes. The Agricultural Marketing Service, on the other hand, has used the average value (farmers' total receipts from or total expenditures on a category of goods divided by the count of units) to construct price relatives. The export and import price indexes are also based on unit-value relatives to a more limited extent.

The Committee believes that in principle the specification method of pricing is the appropriate method for price indexes. The changing unit values of a broad class of goods (say shirts or automobiles) reflect both the changes in prices of comparable items and the shifting composition of lower and higher quality items. The greater rise of automobile prices in the farm index of prices paid than in the CPI reflects the shift toward automatic transmissions and power braking and steering as well as any rise in the price of comparable automobiles. The basic logic of the fixed-weights base index requires that the "pure" price effect be isolated or the index will not measure the changing cost of this fixed weight base.

A special complication arises with the Farm Indexes (and also the export-import indexes) out of the desire to have a unit-value index which can be multiplied by an output index to obtain total receipts or total expenditures. Even if this is a legitimate condition to put on price indexes (a question we need not discuss), it is, however, neither necessary nor sufficient to construct unit-value indexes. It is not necessary because an adequate set of specification prices, joined to appropriately detailed output data, will yield total receipts or expenditures. It is not sufficient because unless one knows the behavior of prices and outputs of individual qualities of goods, it is usually impossible to construct appropriate average values which have the desired property.¹ Staff Paper No. 11 sheds light on this difficulty.

But the Committee also believes that the true difference between unit-value and specification pricing is one of degree rather than of

¹ Actually the unit values are often modes rather than means: the prices asked for are those of (unspecified) qualities bought in the greatest volume by farmers.

kind, and neither the BLS nor the AMS uses either method of pricing exclusively or in the most extreme possible form.² Thus the CPI uses average values for used cars, houses, and other categories, and the AMS has almost continuously refined the categories within which unit values are computed, for example by dividing a single category of refrigerators into four size groups. Yet both agencies depart from what we believe is the best practice.

In the case of the Farm Indexes, the classes over which unit values are computed are still often too wide. Automobiles are classified only into 6 and 8 cylinder (with prices collected for the lower priced brands); there is only one category each of men's wool suits and living room suites; tractors and crawlers are classified only by horsepower; etc. There is too much room, within these categories, for both misunderstanding by price reporters (who report by mail) and large shifts in the nature of goods priced over time.

It is less obvious that one can err in the direction of excessively fine specification, but we are convinced that the BLS has done so. Our discussion of the problems created by present BLS specifications applies primarily, though not exclusively, to the pricing of clothing, home furnishings, and appliances. It is true that the finer the specification, the more comparable the prices and therefore the better the measurement of "pure" price change. But the immense costs of extreme and geographically uniform specification are ignored by this view:

1. If uniform specifications are prepared in Washington for products which vary greatly over the United States, they must fail to represent large amounts of consumer expenditures.

2. Because of the centralization of specification writing, and the endless task of revising specifications as products change, the specifications lag behind changes in the market. Because it requires additional efforts for a busy field agent to obtain waivers or changes of specification, items are sometimes priced until prices are no longer available (which may be long after the commodity has dwindled to unimportance).

3. If the specifications are very strict, only a very few price quotations can be obtained from a given amount of field enumeration. Price changes are often associated with product changes, or with changes in the retail establishments patronized by the index population, which over time increasingly depart from narrow specifications.

The collection of prices for commodities of detailed specification results in a kind of stratification of reporters that changes over time. Because of the great variety of goods in the market and the frequent changes in styles, fashions, and models, the commodities described by detailed specifications cannot be found in all establishments at a given time or in the same establishments in successive periods. In order to provide for a sufficient number of price reports, the specification has to be generalized in various ways, but even considerable tolerances in the definition do not assure that articles "meeting specification" can be found in the same places at successive dates. The choice of the specification leads to a selection of outlets that cannot be described in explicit terms, but it is almost certain that these outlets become less and less representative over time. Specifications must be changed constantly in order to provide price statistics that are in

² See Staff Paper No. 2 for illustrative comparisons of prices collected by the two agencies.

some sense representative of the market. The effect of changing specifications on the index depends on the calculation procedures and involves the measurement of quality differences. When the calculations of price relatives are based on the prices of commodities with different specifications, the precision of the specifications has served only as a control on the agents' judgments on the equivalence of commodities at different dates.

Specifications are soon outdated and must be changed so frequently that their advantages in the recording of price statistics are offset by the difficulty, even impossibility, of assessing the effect of changes in the number of reporters and alterations in the details of the specification. A large part of the resources used in writing specifications, editing and reviewing the price data for errors in recording, and interpretation could be used in designing more general rules for the recognition of comparable qualities within the stocks of individual establishments, developing methods for the recording of qualitative characteristics of commodities and services, and analyzing the relation between the quality variations and price changes that take place over time.³

We do not conclude that specification pricing should be abandoned, but rather that it be reconstructed along more flexible lines. The basic goal of comparable price relatives can be achieved without imposing uniform, highly detailed specifications on the entire price collection area. Specifications should be set centrally (on a basis discussed in the section on sampling, III, 5), but they should not be "specifications in detail"; that is, men's dress socks or 10- 12-ft. refrigerators may be set centrally, but the field agents should be free to select those precise qualities for which they can obtain continuous and comparable price quotations because the commodities are continuously sold in the outlet. The precise specifications should be reported along with the prices, to provide a continuous and up-to-date stream of information on commodity appearances and disappearances (and thus assist in earlier revisions of central specifications). The actual cost of collecting comparable and also more representative price quotations will be substantially reduced by this controlled decentralization.

Because the procedures recommended in the preceding paragraph represent a substantial change from present procedures, it seems desirable to discuss them in some detail. We envisage specifications encompassing a broader range of qualities or varieties of a product than is now priced, but setting forth the features of the product that are to be held constant from one period to the next. For example, instead of requiring that a gas range be white porcelain, the specification would ask the agent to state whether the range being priced is white or in color and instruct her not to compare a colored range this month with a white range last month unless the white range was no longer sold in this outlet in appreciable volume. When a substitution was made, the

³ A change to less detailed specifications would make possible the more frequent and complete publication of the specifications used and of major changes in them, a step that would greatly increase the usefulness of the published item indexes. At present the user of an item index can get information on the nature of the item to which it refers only by corresponding with the BLS, and if the information requested is not current, it can be supplied only with great difficulty. The published item index on men's socks, for example, refers at times to work socks and at other times to dress socks, but there is no way at all for the user to determine this from the published material.

decision could be made in the processing of the field report whether to make a direct comparison, to make a comparison with a price adjustment for the change from white to color, or to introduce the colored range by linking. We believe that such procedures would continue to give commodity specialists adequate control of the pricing process. We also envisage the elimination of the specification of features that cannot be recognized by a well-trained agent in the field—for example, the number of pair of pajamas cut from a given length of cloth—and of features that are not relevant to the performance of the commodity or the satisfaction derived from it—for example, the exterior dimensions of a table model radio.

Uniform national specifications will still be needed for occasional studies of geographical price differences. Such studies, however, do not contribute directly to the production of the CPI, and the quality of CPI should not be impaired to facilitate them.

3. QUALITY CHANGES

If a poll were taken of professional economists and statisticians, in all probability they would designate (and by a wide majority) the failure of the price indexes to take full account of quality changes as the most important defect in these indexes. And by almost as large a majority, they would believe that this failure introduces a systematic upward bias in the price indexes—that quality changes have on average been quality improvements.

We have very little evidence at our disposal with which to support—or deny—the belief in progressive quality improvement. Indeed we are impressed with how little empirical work has been done on so widely held a view and potentially so important a problem. Even the concept of quality change is not free of difficulty. Changes in buyers' tastes will lead to the appearance of new goods—an uncontroversial example would be fashionable apparel—which are not improvements judged by either previous or subsequent tastes, and the line separating taste changes from quality improvements will depend on the time span invoked.

Great as the difficulties are, however, we think it is possible to go beyond the recommendation that more research be done on the problem.

One form of the quality problem is the appearance of new goods: television, blankets made of synthetic fibers, new drugs. We believe that new products can and should be introduced into the indexes much more promptly than they are at present, and the discussions of specification pricing (III,2), new goods (III,4), and sampling (III,5) indicate methods which we believe are operational. Staff Paper No. 2 should also be consulted on this range of problems. The procedures we recommend will not take full account of new products, but will serve to reduce greatly a lag that is now too large.

But the main quality problem will remain: how should one deal with the steady advance of medical knowledge; the annual appearance of new models of consumer durable goods; the appearance of new kinds of retailers; etc.? In general there is no known method of coping with these problems on a current basis, and the current price

indexes must ignore them.⁴ We are not so pessimistic, however, about beginning to treat of quality changes in the annual price indexes which we shall recommend as supplements to the current program. We believe that the following lines of attack deserve intensive and persistent experimentation.

(1) Often there is an array of varieties of a commodity available at a given time, such as automobiles or refrigerators. It is then possible to study the effects of (e.g.) weight, horsepower, and other characteristics upon the price of the commodity and thus to deduce implicit prices of these characteristics. The succeeding year, when the entire spectrum of characteristics has changed (for example, all cars have more horsepower), the implicit price relative for increased horsepower can be deducted from whatever price change actually occurred.

This method has been employed by several students of the quality problem⁵ and is potentially of wide applicability. As an instance of its nature, Griliches has calculated the relationship between the prices and characteristics of new automobiles; his regression equation for 1950 prices is—

$$\log P = .365H + .111W + .192L - .054V_1 - \text{constant}$$

where

P is list price

H is brake horsepower (in 100 h.p.)

W is shipping weight (in 1,000 lb.)

L is overall length (in 10 inches)

V is a dummy variable $V=8$ engine ($V=1$) or 6-cylinder engine ($V=0$).

If this system of implicit prices for 1950 is applied to an index of automobile prices of the "low-priced three" cars, one finds that the corrected average price of an automobile fell by 18 percent from 1950 to 1959. The CPI reports an increase of 31 percent, and would have reported an even larger increase if list prices (which ignore all changes in dealers' concessions) had been employed since the retail automobile market was weaker in 1959 than in 1950. This method of estimating quality change deserves extensive exploration and application.

(2) Technological characteristics of products offer a second avenue to the estimation of quality changes. The technical characteristics are usually diverse, and hence impossible to sum, but often a single characteristic may be of special importance to the buyer. Use of this dominant characteristic as the measure of quality will be imperfect, but less so than complete disregard of quality change. Some instances of this method may be found in the consumer goods area: For example, if the average duration of a hospital stay for an appendectomy has fallen by half over a period, then the effective cost of the hospital service should be halved relative to the cost of a hospital stay of fixed duration.

⁴ Richard Ruggles believes that the current practice of ignoring quality improvement and new products (see sec. IV) is arbitrarily assuming these elements to be zero in their effect and that there would be an increase in accuracy of the price index if some other more reasonable but equally arbitrary allowance were made for these elements. He believes further that an arbitrary allowance for quality change and new products would put the user of the price index on warning as to the existence of such an arbitrary element in price measurements.

⁵ Among them Andrew T. Court, "Hedonic Price Indexes," in *The Dynamics of Automobile Demand*, New York, 1939; Richard Stone, *Quantity and Price Indexes in National Accounts*, OEEC, Paris, 1956; and Zvi Griliches (Staff Paper No. 3).

(3) The collection and analysis of consumer appraisals of comparative qualities by means of attitudinal surveys deserve exploration. The collaboration of psychologists, technologists, and survey specialists could be brought to bear upon the question whether such surveys yield stable and meaningful results.

The foregoing discussion refers to consumer price indexes, where the cost to consumers of maintaining equivalent satisfactions provides the criterion for quality adjustment. In the case of producers' goods two alternative criteria are possible. One would make the criterion for determining equal quality the equivalence of ability to contribute to production, as indicated, for example, by the same operating costs per unit of output for a machine, or the same quantity required per unit of final product in the case of a raw material. The other, with which present practice in the wholesale price index in general seeks to conform, finds equivalence of quality in equal production cost. Choice between the two standards depends on the purpose for which the indexes are to be used, and feasibility of application, and the Committee is not prepared to take a stand on this issue. We do believe that better implementation of the present standard requires additional work, especially in the area of machinery and equipment, and that this would permit more complete coverage of prices in this difficult area. We also urge studies of methods of adjustment that would conform to the alternative criterion of an equal ability to contribute to production.

4. THE TREATMENT OF NEW PRODUCTS

New products and services are constantly being introduced into the marketplace, as a result of new technology, changes in consumer tastes, and the rise of incomes. For every successful innovation, there are many that fail. The successes, however, may quickly win a major share of the market or completely displace old products. Some new products—television sets, for example—are radically different from anything previously available. Most new products, on the other hand, are simply new varieties of older products—for example, nylon socks or filter cigarettes.

The treatment of new products presents a serious problem for any price index. An attempt to introduce all innovations into an index as soon as they appear would clutter the index with the failures that never attain appreciable importance. On the other hand, if new products are introduced only when the old items are completely displaced, the index will become seriously obsolete and will fail to reflect the price movements of the "volume sellers" much of the time.

Inadequate recognition of new products can create systematic bias in a price index because there is a typical price history of a product over its life in the market. New products are usually introduced at relatively high prices and their prices fall as they gain acceptance, owing to the economies of producing them on a larger scale and to improvements in the technique of production that come with time and experience. The price of a mature product or service is likely to be at the lowest level in its history relative to other prices. Finally, in the "old age" of a product, its relative price will often tend to rise as the scale of production contracts and economies of scale are reduced. The recent history of local transit fares illustrates this last phase. (For commodities, two additional phases sometimes occur after production

has ceased—the falling price of remainders and the rising price of antiques.)

The possibility of bias can be discussed in terms of this typical life history if we take as a point of reference an index that includes new, mature, and disappearing products in proportion to their importance in the market or in the expenditures of the index population. Relative to this standard; an index that overrepresents new products will tend to be biased downward and an index that underrepresents them will tend to be biased upward. This upward bias is distinct from, and for some purposes additive to, the result of failure to reflect the fact that at its high initial price a new commodity may be considered a better value by some buyers than the item it replaces.

Although in principle the treatment of new products may err in either direction, we believe that in practice the problem is decisively that of introducing new products too late and retaining old products too long. In the CPI, for example, new automobiles were not introduced until 1940, used automobiles (which workers then usually bought!) until 1952; refrigerators were introduced only in 1934; etc. Another way of documenting the delay is to notice that no instance of the premature inclusion of an unsuccessful new product is known to us.

The tendency for indexes to lag behind changes in the market arises for two different reasons. The less important is the way in which the concept of the Laspeyres or fixed-basket index is sometimes interpreted—so that substitutions or additions of items between major revisions are considered inherently undesirable and to be avoided if at all possible. We do not interpret the concept of the fixed basket in this way. Although it seems desirable to keep the weights for classes of commodities or expenditures fixed between revisions of an index, within these reasonably broad classes the compilers of the index should be free to add or delete items or to reassign weights so as to take account of the appearance and disappearance of products. This is in fact done in practice, and the main problem is to make the necessity for it understood by users of the indexes.

The more serious source of lag lies in the slowness with which knowledge of changes in the nature of goods on the market reaches the people who specify the items to be included in the indexes. Our observations suggest that in many cases the need to introduce a new variety of a commodity is felt only when it becomes impossible to price the old variety because it has disappeared from many outlets. In this case the BLS agent or the respondent to a mail questionnaire will report that it is not possible to provide quotations for the item specified, and a mounting pressure of such reports will lead to the specification of a substitute item. So long as it is possible to provide quotations for the disappearing items they will tend to be provided, even though the item is no longer the volume seller, because it is much simpler for the agent or the respondent to give the information requested without comment than to explain the offering of alternative data. Where the disappearance of an item is slow, or where two varieties coexist for long periods, this can result in very substantial lags in the introduction of new items into the index.

The Committee feels that the agencies constructing price indexes should give high priority to the exploration of methods of introduc-

ing new items into the indexes at an earlier date. Our recommendations with respect to sampling procedures (III, 5) will make some contribution to this desired end. In the case of the CPI, the agents of the BLS undoubtedly can anticipate many changes in items and specifications well before they are actually made. Under present procedures, they seldom have an opportunity to act on this knowledge until they are unable to secure quotations on the old item. One aspect of the problem is to find a way of marshaling the information on new products that is constantly available in the field and bringing it to bear on the problem of specifications.

A part of the problem of new varieties will take care of itself if the BLS accepts the Committee's recommendations on the relaxation of specifications for the CPI (III, 2). In effect we recommend permitting the price of a considerably wider range of varieties of an item from outlet to outlet and from place to place so long as the identity of the item priced is held unchanged in the period-to-period comparison in a given outlet. This would permit an agent to introduce new varieties as soon as they become volume leaders in a given outlet. The question of whether a new variety of a commodity is purchased by families in the index population would be handled in the sampling of outlets. The introduction of completely new items would have to be based on decisions made centrally as at present.

5. PRICE INDEXES AND SAMPLING

The data used in computing the value of a price index are ordinarily derived almost entirely from a highly complex network of samples—samples of goods and services, samples of localities in which prices are collected, samples of actual price reporters, and samples of points in time. It is therefore apparent that a value of an index depends upon the particular samples from which the basic data are obtained, and that different samples will lead to possibly different values of the index. Under such circumstances it is natural to ask how far an observed value can be expected to deviate from a "complete populations" index. Assuming that sampling procedures are employed which provide an unbiased, or "nearly" unbiased, estimate of the "complete populations" index, the customary answer would be specified in numerical terms as the sampling precision of the index. Such measures of precision are not available, at least in published form, for any of the currently prepared price indexes.

All producers of index numbers continually reaffirm their belief in the need for "good" data and hence for the use of "decent" sampling procedures. Furthermore, they state categorically that probability model sampling procedures should be employed whenever possible. They also, however, set forth three related arguments for not attaching measures of sampling precision to the end product. These are:

1. The Laspeyres fixed-base index, upon which most price indexes are modeled, requires that the prices of a sample, or fixed market basket, of goods and services be followed through time. Since the universe of commodities available to the consumer is continually changing, it is necessary to make a variety of adjustments in the sample and in observed prices to account for the disappearance of

old items, for the changing quality of continuing items, and for the appearance of new items. There presently exists no well-defined, "best" set of adjustment procedures and so the index is subject to a *procedural error*, which exists even though all commodities, localities, and price reporters are used in preparing the index. It is then argued that the sampling error is probably small in relation to the procedural error and that it is therefore neither necessary nor desirable to attempt to obtain estimates of its magnitude.

2. Because a complex set of adjustment procedures is required in order to follow the prices of an original sample of goods and services through time, it is frequently stated that it is actually impossible to define or estimate that portion of the sampling variability of an index that arises from the sampling of commodities. Hence it is impossible to define or estimate the sampling precision of the index itself.

3. A third argument admits that it might be possible to employ probability sampling for all components of a price index, and by inference therefore admits the possibility of computing a measure of sampling precision. But the great complexity of the design and data gathering operations for an index are then stressed and the conclusion is reached that the attainment of this goal would require the use of more or less unlimited resources.

The Committee accepts no one of these arguments in its entirety, and feels that it is both possible and necessary to estimate and publish the sampling precision of a price index.

The argument that it is impossible to discuss sampling precision because of the changing nature of the universe of commodities is clearly basic to a consideration of the other two arguments. This can be refuted as follows. Assume a set of adjustment procedures, designated by P , which are used to follow a sample of goods and services through time, where sampling variability arises only from the fact that a sample of items is selected at time zero. (The determination of the details of such a set of procedures is the concern of much of this report. If individuals and organizations cannot agree on at least the major outlines of a reasonable set of procedures, index numbers should not be computed with the present Laspeyres-oriented approach.) Then some well-defined sampling procedure will be used to select a sample of items from the universe of such items as it exists at time zero. If one now thinks of drawing an indefinitely large number of independent samples in accordance with the defined sampling procedure, and of following each of these through to time t in accordance with P , the resulting values of the index will define the sampling distribution of the index with respect to the sampling of items. The variance of this distribution is a perfectly acceptable measure of sampling precision for the index, and it includes a component for any non-uniqueness which may exist in P . Furthermore, an estimate of this variance can easily be obtained by the simple expedient of actually drawing two or more independent samples of items—i.e., through the use of *replicated samples*. It should be observed that the use of two independent samples, for example, does not mean that each sample must be as large as the desired overall sample of commodities. Each sample may be only half as large as the overall sample, and the published index would be the average of the two resulting indexes. Of course, the reliability of the estimate

of variance would improve as the number of independent samples increases.

Assuming that it is both conceptually and economically possible to estimate the precision of a price index due to the sampling of commodities, we next consider the argument that this precision is dominated by the procedural error and can therefore be ignored. Some investigations reported in Staff Paper No. 4 suggest that the procedural error of current consumer price indexes may indeed dominate the sampling error,⁶ for a value of the index some three years after the base period, although empirical investigations of the overall effect of procedural error are almost as lacking as those of sampling error. In the opinion of the Committee, this does not mean that sampling error can be ignored. In particular:

1. If the goal is to estimate the *level* of the "true" index at various points in time and if resources are fixed, then the most efficient way of improving the accuracy of these estimates would be to divert resources from the maintenance of a relatively large sample of commodities and to use these resources in basic research aimed at reducing the magnitude of the procedural error. It is clear that good estimates of sampling precision and of bounds on the procedural error are required in order to make judgments of this kind.

2. If the goal is to estimate short-term *changes* in the level of the "true" index, then it appears likely that sampling error will be more important than procedural error and hence an estimate of sampling error becomes essential.

3. The construction of a price index involves not only a set of adjustment procedures and the sampling of commodities but also the sampling of localities and the sampling of price reporters within these localities. Just as there must be a balance between the procedural error of the index and the error due to the sampling of commodities, so also must there be a balance between these errors and the sampling errors arising from the other parts of the design. Again it is impossible to discuss such a balancing operation unless some attempt is made to measure these components of error.

Not only are estimates of error for the various components of a price index needed for internal design purposes, but they must also be available in published form to assist those who wish to use the indexes in a critical fashion. Such publications should give precise descriptions of the methods used in drawing samples and of the formulas employed in estimating the population index. The publications of the agencies currently producing index numbers are too much oriented toward the general reader, and additional details must be presented for the benefit of the scientific community.

Although the Committee recommends that every effort should be made to use some appropriate form of probability sampling in the selection of each sample that enters an index design, particularly to guard against nonmeasurable biases from sampling and estimation, it recognizes that the sampling of goods and services poses an especially difficult problem. There are, however, convincing reasons for at-

⁶ This observation relates to an all-item consumer price index. If attention were focused on subgroup indexes, then there might well be instances in which the sampling error would dominate the procedural error. Improvements in the accuracy of such subgroup indexes could then be obtained only by increasing the number of commodities drawn from particular subgroups. Considerations of this kind would increase the number of item in the total index over what is implied in the following discussion of the all-item index.

tempting to use probability sampling methods, at least as a guide, in the original selection of items. Some of these reasons are: (a) The replicated sample approach can provide an estimate of sampling precision for almost any type of sampling procedure, provided only that the procedure is defined in such a manner that repeated and independent drawings can be made, but it cannot even indicate the existence of bias. The only way to ensure that biases due to sampling and estimation are small or nonexistent is to use appropriate probability sampling methods. (b) A probability model will make clear the manner in which one can obtain two or more independent samples of goods and services. (c) Even the mere attempt to make the sampling of goods and services conform to some appropriate probability model will force one to think explicitly and to make decisions about problems of definition and estimation which exist no matter how such a sample is chosen, but which can too easily be ignored with judgment procedures.

The exact details of a method for selecting a probability sample from the universe of goods and services as it exists in the base period can only be obtained through careful investigation. However, the general format that possible procedures would probably follow can be indicated, using the Consumer Price Index by way of illustration. Items of expenditure, as determined, say, from Consumer Expenditure Surveys, are divided into major groups (food, housing, apparel, and the like), then into subgroups, sub-subgroups, and so on. Ultimately, this subdivision process leads to what may be termed specific items—e.g., oranges, frozen peas, radios, men's sport shirts, women's nylon stockings, auto registration, and men's haircuts. Presumably one would be able to obtain an almost complete enumeration of the universe of specific items and their associated base year expenditure weights. These specific items can then be grouped into strata, using any available information about substitutability, similarity of price movements, and other related variables. The *first* sampling operation would then consist of selecting one or more specific items out of each stratum. If this selection were made with probabilities proportionate to base year expenditure weights, an unweighted average of the price relatives of the sample specific items would provide an estimate of the weighted average of specific item price relatives for the entire stratum. (Banerjee and Adelman have discussed the details of related procedures, and references to their papers are given in Staff Paper No. 4.)

When one draws a specific item into the cluster, he has in most cases actually drawn an entire *cluster* of *specified-in-detail* items into the sample. Thus there are different varieties of oranges, men's sport shirts made from different material and differing in quality, and so on. The current procedure used in the Consumer Price Index calls for the selection of one or more specified-in-detail items from the cluster defined by each of the selected specific items, and this is the *second* sampling operation to be considered. The chosen specified-in-detail items are the ones on which price quotations are to be obtained. (Occasionally a specified item will also be a specified-in-detail item, but this will be the exception rather than the rule.) At this second level of sampling, the problems become much more difficult than at the first level. Complete lists of specified-in-detail items will be difficult, if not impossible, to obtain; some specified-in-detail items may not be purchased by the consumer group to which the index is supposed to refer;

and expenditure weights may not be available for many of these items. Possibly anything that one can do at this level (e.g., using a restricted list of specified-in-detail items instead of a complete list, or assuming equal base year expenditure weights when the actual weights are unequal) is going to be only an approximation to what one would like to do, but at least this type of approach can be described accurately and it should be possible to investigate the effects of some of the approximations that are used.

Even though it were possible to use probability sampling for all components of a price index, the Committee recognizes that "simple" estimates of error are required for most purposes. The necessity for designing a complex sampling operation so that "simple" estimates of error can be obtained has long been recognized. This need becomes overwhelming in the case of the price index where the number of commodities entering the index is large and where the quality adjustment procedure makes it difficult or impossible to apply variance estimating techniques derived from sampling theory to all components of the design. Furthermore, these estimates of error for a price index have to be made more or less continuously since the sampling errors can be expected to increase with the length of time from the base period. Some suggestions concerning this type of design for the Consumer Price Index are given in section VII of Staff Paper No. 4. Basically, "simple" estimates of error can be obtained through some form of replication such as was mentioned under commodity sampling.

Sampling considerations also suggest two possibilities for major changes in price index construction which should be given serious attention in the future. They are:

1. Index numbers of the Laspeyres type are based upon market baskets of commodities which remain essentially unchanged between major weight revisions, except for adjustments which are made either to account for the changing quality of items or to recognize marketplace substitution of "new" items for "old" items. Some individuals have advocated drawing a completely new sample of items at fixed intervals from the universe as it exists at the time of drawing, together with a chain approach for obtaining comparisons over longer periods of time, but this suggestion is not likely to be adopted in its entirety. It should, however, be possible to effect a compromise between these two extremes and thereby gain some of the advantages of each. Thus one could set up a rotation schedule so that each item remains in the index for some fixed period of time, say 1 year, or 2 years, or 3 years, and so that a fixed fraction of the items are replaced each month, or quarter, or year, by newly selected items. This type of approach would obviously require some departures from the strict Laspeyres concept of an index number, but then these departures are also required by currently used quality and new item adjustment procedures.

2. The Consumer Price Index is basically city-oriented. That is, indexes are computed for each city in the city sample, and these indexes are weighted to obtain the U.S. index. This emphasis on city indexes does not appear to be the most efficient way of obtaining the U.S. index. If one views the index in terms of U.S. average weights and average prices, then it is clear that quite a different sample should be used, for example, to obtain a "good" estimate of the average price of a newspaper than would be used to obtain a "good" estimate of

the average price of a used car or of a woman's coat. In other words, the size of the "best" city sample for an item depends upon the cost of obtaining a price quotation and upon the variability of the item's price from city to city, and thus the size of the "best" city sample will differ considerably from item to item. (This fact is recognized in the present CPI design through the use of different monthly pricing cycles for different items.) It is of course true that aggregation according to a Laspeyres index calls for price quotations to be weighted in proportion to population and to value, and that a complete set of value weights could not possibly be obtained for all cities in which one would be able, for example, to collect newspaper prices. But this difficulty might be overcome, for example, by deriving the Consumer Expenditure weights for the population of cities in a region rather than for a number of individual cities in the region. An added benefit of such a change in emphasis might well be that it would become more feasible for BLS to employ selected data from other sources in the index computations, e.g., from the Monthly Retail Trade Report of the Bureau of the Census.

6. SEASONALITY AND SEASONAL ADJUSTMENT

Each of the price indexes under review covers many commodities that are subject to substantial seasonal fluctuations in both prices and quantities consumed or sold. Seasonal changes in quantities are often, but not always, associated with seasonal changes in prices.

Amplitudes and patterns of seasonal variation differ greatly among the individual components and groups of items within each of the indexes. The seasonal movements of quantities are similarly diversified. Intra-annual fluctuations in both quantities and prices are particularly strong among the commodities in the CPI food group and the WPI farm products and processed food groups. Fresh fruits and vegetables have much larger seasonal amplitudes than other items, but meats, poultry, eggs, and milk, to name only some of the more important commodities, all show marked seasonalities. Substantial seasonal influences are, of course, also at work in apparel with its new spring and fall lines and summer and post-Christmas sales. For the major consumer durables, especially automobiles, model changes are very important. The timing of the seasonal peaks and troughs, expansions and contractions, varies greatly among the component price series, however, so that these movements, as now measured, offset each other to a large extent, leaving only relatively small seasonal changes in the index as a whole. (This applies to each of the price indexes reviewed.) Nevertheless, seasonal influences may and at certain times do dominate the behavior in the very short run of a comprehensive measure of average price changes such as the Consumer Price Index. The overall sensitivity to seasonal factors of the Wholesale Price Index seems to exceed somewhat that of the CPI, while the Indexes of Prices Received and Paid by Farmers (especially the latter) appear to be less subject to such influences.

All the major price indexes now in use are, in fact, a cross between unadjusted and seasonally adjusted indexes, although they are published as unadjusted series. They employ annual rather than monthly weights, which is roughly equivalent (to the extent that interseasonal movements in the base year are only seasonal in character) to using

seasonally adjusted monthly weights. Prices of commodities that are traded in significant quantities throughout the year represent unadjusted prices, as do the available market prices of commodities that disappear during part of the year.

Price quotations for these latter ("disappearing") commodities—those which are not sold in certain months of the year or for which prices are not available in certain months even though some trade in them does take place—are either held constant at the last quotation in the off-season months, or extrapolated from the date of the last quotation by the index of supposedly related items. By either method, when they return to the market the actual quotation is introduced, and an abrupt break may appear in the series. The former method is used generally in construction of Indexes of Prices Received and Paid by Farmers, and for disappearing apparel items in the WPI. Until April 1959 it was also used for farm products and foods in the WPI. Since that date, prices of these items in the off-season have been extrapolated by the movement of the product class in which they fall. In the CPI, prices of several fresh fruits have been extrapolated in the off-season by the group index for total fresh fruit prices, and those of seasonal items of apparel by prices of year-round apparel items.

The problem of seasonally disappearing commodities cannot be avoided by seasonal adjustments, but the difficulties are less than in an unadjusted series: Monthly price fluctuations are much smaller, and movements of related commodities on which imputation may be based from the date before disappearance to the date of reappearance presumably are less dissimilar when differences in their seasonal patterns are eliminated. With adoption of a policy of annual revision (see III, 8), final estimates for disappearing commodities should be based on interpolation between the dates of disappearance and reappearance rather than on extrapolation from the former date. This will avoid a "break" in the revised series at the time the commodity reappears.

The Committee finds that the major purposes for which the price indexes are used—examination of cyclical and longer term price movements, wage and price escalation, and deflation of the national product and other important value series—are best met by seasonally adjusted indexes. This finding is hardly surprising, since it has already proven to be the case with other key economic series. We recommend that the responsible agencies prepare and publish seasonally adjusted series. Both the aggregate indexes and major subgroups should be presented on this basis.⁷ The publication of the "unadjusted" price series should be continued on the present basis. These "unadjusted" measures, despite their ambiguities, provide information of value to users interested in seasonal price fluctuations.

There is no ideal method of dealing with the problem of seasonality of quantities: in the limiting case in which the budget of each season

⁷ For certain commodities price changes are customarily made annually at the date of model change. In addition there may be a seasonal pattern superimposed that reflects varying discounts from list prices. When price movements have been predominately in one direction (recently upward) over a period of years, care must be taken to isolate the two types of price changes to prevent the seasonally adjusted price series from smoothing the annual price change over the year rather than appearing in the month in which list prices are changed. This point becomes especially important if the direction of annual price change should be reversed. Similar situations are met in many other series; the public education payroll component of personal income, in which major changes in level occur at the start of each school year, is a typical example.

was unique no interseason comparison of prices would be meaningful. But the most attractive theoretical solution lies in the direction of seasonal weights (see Staff Paper No. 5). With (say) four sets of seasonal weights, the comparisons from year-to-year would be unambiguous. The comparison of prices among seasons would not be so simple, but it would probably be possible to devise, through the data on consumption of nonseasonal goods, budgets with approximately equal levels of utility for the various seasons. We do not recommend adoption of seasonal weights for the current indexes, in light of the complications and costs that would be entailed, but in areas where seasonal weights differ greatly, this type of solution deserves serious consideration.

7. CONSUMER DURABLE GOODS

The commodities which yield services over a substantial period of time—houses, automobiles, television sets, furniture—offer special problems for consumer price indexes. In principle, they may be dealt with by either of two general approaches.

The first is to include in the weights the purchases of the durable goods during the weight base period, and to price the goods currently. In addition, there will be separate weights for current operating costs of the durable goods possessed by members of the group covered by the index. The second method is to ignore purchases of durable goods, and instead to measure the cost of the use of the goods. The use cost in the base year will be the weight. The price index may be either rents charged for the use of similar goods or the prices of the components of this use cost.

Both the BLS and the AMS use the former, or purchase, method of dealing with durable goods. The method, however, encounters a number of problems.

1. The rate of purchase of assets fluctuates greatly in response to cyclical fluctuations in income and other factors, so purchases in a given year may be very abnormal. To avoid distortion, weights should be based on an average of several years. In its last revision of the CPI, the BLS used the experience of the entire 1940-50 period for housing (see 3 below).

2. The rate of purchases depends not only on the consuming habits of the population but also (and this is not true of other categories) on the rate of growth of the number of families, migration, the rate of shift from tenancy to ownership, etc. Index weights for individual cities in which there is a high rate of net immigration are thus larger for housing than they would be in cities with a stable population. The weights do not represent the expenditure pattern of those continuously living in the city but instead are weighted averages of such families and those recently immigrating. This is a tenable concept of the "average" family but it implies a corresponding standardization of rates of change of the population characteristic.

3. The proper weight for each expenditure category is net expenditure, after deduction of sales or trade-in of used assets. This is most important for automobiles and houses. For automobiles, BLS obtains net expenditures directly from the consumer surveys. For houses, duplication was eliminated by the following technique when the present weights were established. Home owners were asked when

they bought their present house. From their answers, an average annual rate of purchase in the 1940-50 period was computed, in which owners who bought more than one house during the period were counted only once. The weight for home purchases in each city is the estimated 1951 market value of homes times the annual rate of acquisition among home owners applicable to the city times the proportion of owners in the city.

4. The proper division between new and used items of the total weight assigned to each commodity should reflect the net purchases of each by the index group from other groups in the society. In a price series covering the entire population there would be a seller for each buyer and none of the weight would attach to used items since the group as a whole would have no expenditure (aside from transfer costs which should be weighted and priced). Where the index covers only part of the population this equality need not exist, and used items may receive some weight. However, the BLS procedure, utilizing gross purchases less trade-ins, assigns too high a weight to used cars and houses. (Thus 2.0 of the 4.8 percent assigned to automobiles is assigned to used cars.)

5. Use of most durable goods requires expenditures for fuel or power, replacement parts, and repairs. Changes in such requirements accompanying technical changes in the durable goods themselves must, under the expenditure method, be regarded as representing a quality change in the durable goods. However, no technique has been developed to adjust for such quality change.

6. When consumers pay interest on loans a question arises both as to the weight to be attached to the interest payment and to the price series to be used to measure the cost of interest. Not all consumer interest payments are due to purchases of durables, but most are. In the case of mortgage interest, the largest component, the BLS has developed ingenious and acceptable techniques for deriving the weight to be attached to interest, and adjusts the price series for interest as the dollar amount of mortgages changes with changes in the price of houses.

For the interest series itself the BLS uses interest rates on new mortgages, rather than average interest rates on all outstanding mortgages. Evidently, the consumer is viewed as "buying" the service of not repaying the debt for the entire period of the mortgage contract, at the time the mortgage contract is made. This Committee considers this treatment as artificial and unacceptable. The costs of borrowing so measured do not reflect changes in the actual costs incurred by the average consumer from month to month or year to year. We recommend substitution of average interest rates on all outstanding mortgages for interest rates on new mortgages in the CPI, and parallel treatment for other types of interest paid by consumers whether or not associated with the purchase of a commodity.

An alternative to the expenditure approach, which eliminates the difficulties cited but introduces others, is to measure expenditures for houses and other durable goods on a use basis, in deriving weights, and to measure changes in actual or opportunity cost of using the goods for the necessary price series. Thus, the Office of Business Economics values the services of owner-occupied houses by estimating the rent that these units would bring if rented. In deflating consumption ex-

penditures, it then utilizes indexes of rents charged on rented units. The main difficulty in the latter approach is that the composition of owner-occupied and rented units is markedly different.

This Committee recommends that BLS investigate the possibility of developing a rent series for units that are actually rented but are as representative as possible of owner-occupied units in structure and location. In some neighborhoods very few units are in fact rented, and those that are frequently are rented furnished, so that development of a representative sample will undoubtedly prove difficult. If such a series can successfully be constructed, it would offer the basis for an improved CPI, and make possible a significant improvement in the national accounts.

We know of no other application of the imputed rental technique except to durable goods for houses. Relevant rental values are not generally available for other durable goods because they are rarely rented or rented under atypical circumstances.

An alternative to the rental-value approach is that described in Staff Paper 6, which proposes an estimation procedure for the components of use cost resting on the principle of maintaining the real value of the consumers' assets.

If a satisfactory rent index for units comparable to those that are owner-occupied can be developed, this Committee recommends its substitution in the CPI for the present series for the prices of new houses and related expenses. We recommend further exploration of the possibility of using the approach described in Staff Paper 6 for possible adoption for other durable goods, and for houses if the rent approach is not successful, but are not prepared at this time to recommend that it be ultimately adopted.

8. REVISION AND CORRECTION POLICIES ⁸

Because of the institutionalizing of the indexes, the Bureau of Labor Statistics has adopted peculiar policies in dealing with the inevitable problem of errors and incomplete information. The BLS does not issue preliminary Consumer Price Indexes: errors are corrected in the index for the month in which they are discovered, unless the error in a major component or in a city or U.S. index exceeds a certain magnitude. Nor does it conventionally revise earlier figures (the "new unit" correction of the housing component for 1942-50 was the major exception). The principle is followed in the WPI, but preliminary and final figures are issued in consecutive months for this index. The AMS has published preliminary and revised figures, and revises backward up to 3 years (aside from major revisions of the entire index) in the light of subsequent data.

It is easy to sympathize with the motives which led to the BLS policies, but not with the policies. The decisive objection to the policies is that they lead to the perpetuation of known errors, and this is simply incompatible with scientific standards. We believe that the agencies can serve the legitimate demands of wage, contract, and parity escalation without being in such a straitjacket.

⁸ In view of the proximate nature of price indexes due to problems of product quality, choice of weights, sampling of prices, etc., and of the consequent absolute size of the probable error of the change in price indexes, Richard Ruggies recommends that the official published price indexes should be reported in terms of full percentage points rather than in tenths of percentage points as is now done.

The preferable procedure is to issue preliminary indexes, which would be the basis of contractual and legislative uses. The final indexes can then be issued with the lag dictated by the flow of information. The present practice of misdating the changes in prices should be abandoned.

Two important instances may be given where retroactive revision is essential in the BLS indexes, at least with present procedures. When seasonal commodities disappear for a time, they are sometimes estimated on the basis of similar commodities' prices until they reappear. Once their prices become available, their previous level should be revised to agree with the new price. A more important but parallel problem arises for prices in cities where collections are made quarterly. For the two months in which no price is available it is imputed on the basis of other cities. Any correction on the next reporting date is absorbed in that month rather than prorated over the nonreporting months.

These desirable retroactive revisions, and those based upon the delayed availability of superior information, can best be made by issuing each year a set of revised monthly indexes for the preceding year.

The price collecting agencies generally rely upon their own price reporting programs for their price data, as indeed they should and must. But occasionally important and reliable outside information (e.g., rents or medical data collected by another government agency) become available. We believe that such information should be fully utilized in the revision of previous indexes.

