

Stanley K. Kulpinski, Stuart J. Cohen, Kathleen G. Perez-Lopez, Bureau of Labor Statistics

### Summary

This paper describes the survey design being implemented for the revision of one of the government's major economic indicators, the former Wholesale Price Index, now known as the Producer Price Index. A brief comparison of the current and revised programs is presented. This is followed by the details of the design of the revised survey, including the development of a sampling frame, probability sampling of items within primary sample units in the field, and identification of suitable estimation formulas. The paper concludes by discussing some of the remaining goals for the revision.

#### I. Comparison of Current and Revised Surveys

It will be helpful to briefly compare the major aspects of the survey currently used for the Producer Price Index or PPI (formerly known as the Wholesale Price Index or WPI) with the goals for its revision. This comparison will be in terms of objectives, sampling methods, and estimation.

The primary objective of the current PPI survey is the production of 2,800 commodity indexes. About 160 industry sector indexes are also produced. In contrast, the first priority of the Producer Price Index Revision (PPIR) will be to produce indexes for each of the 493 Mining and Manufacturing industries, as defined by the 1972 Standard Industrial Classification (SIC) Manual, while the production of about 6,000 commodity indexes is a secondary goal. Production of all indexes for both surveys is (will be) on a monthly basis.

Currently the indexes are based on judgment samples, determined by commodity analysts (professional economists), of large companies known to be producers in the commodity areas to be priced. A random process is not used deliberately in the selection. Similarly, the commodity specifications to be priced (the specific commodities along with their transaction characteristics) from each reporter are determined by the commodity analyst. About 10,000 quotes are collected from about 3,000 reporters. The plans for the revised indexes are to use probability sampling techniques to select both the 35,000 reporters, or primary sample units, and the 165,000 quotes they will be asked to supply. For the revision, we are attempting to use as a primary sample unit a slight modification of the economic concept of a price forming unit (PFU). A PFU is basically the largest set of establishments of a company within which the prices are uniform and for which records are kept centrally. The modification of the PFU is to restrict it to an individual industry (4-digit SIC).

Both the current and revised indexes are modified Laspeyres indexes. However, the current indexes are based on the unweighted mean of company relatives of unweighted average prices for items in the index commodity group, while the revised

indexes will be based on weighted individual price relatives. (A price relative is the ratio of prices from two time periods.) The weights will be based on the sample selection procedure employed and value of shipments data.

#### II. Survey Design of the Revision

##### A) Conceptual Objectives and Principal Elements of the Design

A summary of the statistical goals to be incorporated in this revision is: 1) to use probability sampling; 2) to produce estimates of the error incurred by sampling; and 3) to control or eliminate sources of non-sampling error.

In order to implement a scientific survey design, our first consideration is to obtain a suitable frame. Once a sampling frame is available, we must design and select a probability sample of primary sample units, each consisting of one or more establishments. During initiation of the primary sample units, we then select a probability sample of transactions involving items which we intend to reprice monthly. Our intended method of selecting these second stage units is to use a probability mechanism termed disaggregation. For both those stages of sampling, a reasonable allocation algorithm must be developed.

In the area of estimation, a revised index formula is required. Statistically sound and practically feasible methods of estimating the sampling error must also be derived. Development of non-sampling error estimates will be tackled but undoubtedly will be more difficult.

##### B) Specific Details of the Design

###### 1. Frame Construction Process

The standard frame source for establishment surveys at the Bureau of Labor Statistics (BLS) is the Unemployment Insurance (U.I.) file. It contains approximately 4,000,000 records which encompass almost all establishments in the U.S. economy. It is composed of 51 individual state submissions, each state file the result of a state program designed to collect taxes from employers in order to pay unemployment benefits to individuals. The establishment files are the lists of the employers paying the unemployment insurance tax. A continuing problem that most BLS Programs confront, and the PPIR is no exception, is that the name and address on the file frequently refer to a headquarters unit or tax accountant responsible for forwarding U.I. taxes. Most surveys require the physical location of the establishment so that field representatives can be sent directly to the source of data without going through the additional process of trying to locate the unit through an intermediary.

The PPIR not only has the problems that these other programs have, but there are several other

issues that must be addressed. In order to cover the 493 four-digit SIC Mining and Manufacturing industries, we plan to form 493 independent sampling frames. The frames must incorporate the economic and statistical concepts of the revision, must be mutually exclusive and exhaustive, and of course must be able to be constructed in an operationally feasible manner.

The procedure used presently for frame construction in the PPIR was established and agreed upon by all BLS offices involved in the project. The U.I. file is the basic sampling frame. Basic implies that the U.I. is the foundation upon which modifications are made. These modifications can be segregated into two parts. The first part requires the incorporation of the economic concept of the price forming unit (PFU). The U.I. file currently does not have all the information necessary for the grouping of establishments belonging to the same company. A commercially available file purchased by the BLS contains such additional information. Using both files we form preliminary PFU's. These PFU's contain establishments that are classified in the same 4-digit SIC in either file and belong to the same company. We subsequently match all the establishments from the commercially available file in each PFU with the U.I. file. This requires matching about 50 establishments per industry. The primary purpose for this match is to minimize duplication. To illustrate the problem, assume that a preliminary PFU has been formed using the commercially available file. Suppose this PFU contains three establishments, two of which were easily identified in the U.I. file in the same SIC. The third, however, was located in the U.I. file in a different SIC. After a telephone contact determined that this establishment was classified correctly in the commercially available file, it is moved in the U.I. file from the erroneously assigned SIC to the correct one. Therefore, this establishment appears exactly once in the file and in the correct SIC. The telephone contact is also used to determine each company's record-keeping practices so that PFU's can be made to conform to the data available.

The second type of modification is called frame refinement. Although the creation of PFU's does require some frame refinement, it is only a byproduct of the activity. Another more extensive attempt is made to refine the U.I. file, concentrating on large establishments. It is not considered to be cost effective to carry out this matching beyond the largest 25-30 units because of discrepancies between the U.I. and the commercially available file. Telephone contacts and other sources provide updated information, particularly employment, physical location and SIC classification. These data are then used to replace existing U.I. information.

The final sampling frame, therefore, is a modified U.I. file that satisfies the economic and statistical needs of the revision and whose construction is operationally feasible. (Actually the U.I. file will be used about 95% of the time. We have left open the option of using other sources in other instances where such action would seem appropriate and necessary.)

## 2. Allocation of Sample Units to Industries

Allocation of sample units to a 4-digit SIC is currently a two step process. Based on anticipated budget requests and publication plans, an estimate of the total number of primary sample units that could be supported was provided. A preliminary allocation of sample units to each of the 493 industries is then required in order to apportion the available sample size to the industries to be surveyed initially. Using available Bureau of Census data, an estimate is made of the number of primary sample units required for the industry to insure comparable reliability among the 4-digit SIC indexes and acceptable reliability at summary levels. Several factors are considered in this estimate, primary among which are the number of companies in the industry, the concentration of value of shipments in the largest 20 companies, the number of 7-digit product classes in the SIC, and the total value of shipments for the industry. Presently a simple mathematical model using these inputs is being employed to produce these preliminary allocations.

These are reviewed in the second step of the allocation process. A market study for each industry allows a more precise and detailed look at the industry. The criteria mentioned above are considered but in addition PFU structure, stratification, frame evaluation and expected industry response are also used to adjust the preliminary allocation.

### 3. Stratification

Once the frame units have been determined, they are stratified, explicitly and/or implicitly, based on a number of considerations, which are discussed below.

Homogeneity of Price Movement - We attempt to create strata within which price change among units will be similar, and between which price change differs.

Variability of Price Movement - If there are indications that characteristics can be determined that would yield groups that may have the same mean price movement but would differ substantially as to the variation about that mean, we would stratify according to those characteristics.

Expected Non-Response - Since non-response adjustments will be based upon the explicit strata used in sampling, the expectation of non-response may enter into the decision to use explicit strata rather than implicit strata. That is, in situations where implicit stratification, for either of the reasons described above, may be sufficient to reduce sampling variance, we may opt for explicit stratification of groups of units expected to have different response rates.

For treatment of the three considerations above, we rely on the professional judgment of the industry analysts involved, since currently there is insufficient existing price data to use in more objective analysis.

Publication or Analysis Goals - Although production of industry indexes is the primary goal of the PPIR, an important secondary goal is to produce other indexes, either for publication purposes or for internal analysis. Most prominent among these other indexes are 7-digit product indexes. Therefore, if we have indications that within an industry certain groups of establishments tend to have very restricted production, (i.e. within one or two product classes), we would stratify accordingly. For instance, if we knew that an industry establishment on the West Coast produced one type of product and on the East Coast another type, and if we wished to produce indexes for those products, we would stratify the frame units by region. We would expect to thereby increase the chances of getting sufficient data for both types of products.

Characteristics of PFU's - If the PFU's cross strata lines, or if they are unusually large or distinctive, we would create a separate stratum for them. These units are often all selected with certainty.

Whether to stratify explicitly or implicitly often depends on practical considerations and the strength of the indication for stratification.

#### 4. Allocation to Strata

If value of shipments data is available from the Census Bureau or from some other source for the strata that have been created, the sample units for the industry are generally allocated to them proportionally to those values. This is expected to approximate optimal allocation if the assumption is reasonable that the strata population variance of price movement is directly proportional to the square of the average value of shipments for the strata. When value of shipments data is not available for those strata, groups of strata are formed for which such data is available. These groups are kept as small as possible. Sample units are then allocated to the groups in proportion to their total value of shipments, and to the strata within the groups in proportion to total employment for the strata. We expect that employment is an adequate proxy for value of shipments within the groups.

#### 5. Sample Selection

Within each stratum a systematic selection of the allocated number of sample units is made, with each frame unit having a chance of selection proportional to its total employment. Here again, as in the allocation to the strata, we expect that the employment figures yield good approximations to the proportions we would obtain from the value of shipments, if those data were available.

#### 6. Allocation of Quotes to Sample Units

An algorithm has been developed for the allocation of quotes to the sample units based on the number of 7-digit product classes in the industry and the sizes of the sample units. This allocation is usually from 3 to 8 quotes but can be increased for very large sample units.

#### 7. Selection of Quotes From the Sample Units

During the design of the pilot phase of the revision, (first four industries selected), there was an effort to develop a probability mechanism for sampling transactions at the establishment site. The result was a process adapted from one used in the Consumer Price Index Revision Program; it is known as disaggregation and proceeds as follows. After the field representative has located the proper primary sample unit, he has the respondent classify his establishment's product records into broad categories. The product categories are flexible and based on records availability at that sample unit. Product categories are then selected through systematic sampling, with probability proportional to size. Successive nested classifications follow, and more detailed products are selected. Ultimately the process identifies a specific product, specific in the sense that it should have no other price-determining features which could be used to further distinguish between products.

We then disaggregate to a unique type of buyer and unique terms of sale for the selected product. At each stage we attempt to provide each category a chance of selection proportional to its actual value of shipments. Respondent estimates, ranking and equal probability are first, second, and third alternatives, respectively, to actual value of shipments.

Our intention is to use disaggregation throughout, but at the very least we intend to disaggregate at the highest or most general levels. We believe that any bias incurred due to non-probability sampling at lower levels is much less significant than that due to a higher level departure from disaggregation.

#### C. Estimation Methodology

The general form of the Laspeyres index is the ratio of an estimate of current period revenue using a base period mix and quantity of items to the base period revenue for the same mix and quantity of items. Symbolically, this can be written:

$$I^t = \frac{\sum q_0 p_t}{\sum q_0 p_0}$$

where  $I^t$  is the index for time  $t$ ,  $q_0$  is a base period quantity of the item shipped,  $p_0$  a base period price, and  $p_t$  a price at time  $t$ . The summation is over all items represented by the index. An alternative form uses revenue weights and price relatives:

$$I^t = \frac{\sum (q_0 p_0) p_t / p_0}{\sum q_0 p_0} = \frac{\sum w_0 r^t}{\sum w_0}$$

Here  $w_0 = q_0 p_0$  refers to the base period revenue and  $r^t = p_t / p_0$  is the price relative

that moves the revenue to its estimated current level using base period mix and quantities. In the PPIR, we generally use the latter form since we collect revenue rather than quantity data for selected items.

The actual estimation formulas developed for the industry indexes of this survey are based on the general form, and take into account the following points:

- 1 - we are dealing with a sample, not a universe of transactions;
- 2 - we are concerned with products shipped out of an industry, which in this context is called net output;
- 3 - not all sample units respond;
- 4 - there is universe data available from the Census Bureau that can be used to improve our estimates.

As a result of the first point, a weight must be applied to the revenues of selected items to expand their value to the universe level. This weight is the inverse of the chance of selection through all stages of sampling.

To handle the second point ideally, we would select establishments proportional to their revenue from shipments going out of the industry and disaggregate only from an establishment's products that are shipped out of the industry. Since practical considerations make such a procedure impossible, we are forced to make two assumptions. The first is that within an establishment, the price movement for an item in a 7-digit product does not depend upon whether or not the item is shipped out of the industry. The second is that the proportion of a 7-digit product that is shipped out of an industry is the same for all establishments in the industry. We are able to estimate these proportions, called net output factors, using input-output data from the Bureau of Economic Analysis and the Census Bureau. The first assumption allows us to use price data from establishments' total production in the index; the second allows us to reduce the revenue weights from gross output to net output at the product within industry level rather than at the more detailed item level.

The third point, the existence of non-response, is treated by applying a factor that expands the revenue data supplied by responding sample units to an estimate of the revenue for all sample units. This expansion is based on employment data and is done within the original sampling strata, which may have been collapsed.

The fourth point, the availability of relevant universe data, resulted in the development of several possible index formulas. Each of these incorporates the universe data in a different manner, creating ratio estimates of the base period and current period revenues. Investigation of these alternatives will proceed to determine whether the calculation of  $\gamma$  should be changed.

The resulting general form for the industry indexes of the PPIR is then:

$$I^t = \frac{\sum \gamma E_s \alpha_j a_h w_h r^t}{\sum \gamma E_s \alpha_j a_h w_h}$$

where:

- $r^t$  is the ratio of the price at time t to the price during the base period for item h, which is in 7-digit product j;
- $w_h$  is the revenue from item h shipped during the base period;
- $a_h$  is the weight applied to item h to account for all stages of sampling;
- $\alpha_j$  is the net output factor for 7-digit product j to which h belongs;
- $E_s$  is the non-response adjustment factor for stratum s;
- $\gamma$  is an adjustment factor using universe data;
- $I^t$  is the industry index;

the summation is over all items that are selected from establishments sampled in the industry.

The product indexes of the PPIR are very similar in form to the industry indexes. However, the second point discussed above, the net output aspect of the industry indexes, is not relevant for the product indexes. This eliminates the net output factor,  $\alpha_j$ , from the index formula for products. The adjustment factors using universe data,  $\gamma$ , would be different from the ones used for the industry indexes, and of course, the summation would be over all items classified in the product category instead of an industry.

To date we have not developed the methodology for computing estimates of the sampling variance of our index estimates.

### III. Some Remaining PPIR Goals

#### A. Feedback of Collected Data into the Design

The current design for the PPIR, though statistically valid and an improvement, is quite crude. We do not have enough relevant data to create a sophisticated design. However, our survey design goal for the future is to use the results of the first revision experience to refine the design for the later industry samples.

Two examples of refinements that should be possible to the design of future samples are the better allocation of sample units to industries and the better use of employment data in the design. When we are able to produce estimates of the sampling errors of the PPIR industry indexes, we will be able to evaluate the success of our allocation of primary sample units to the industries. Adjustments can be made to the algorithm if we find that it is not resulting in comparable reliability among the industry indexes and acceptable reliability at summary levels. Examination of initial employment figures versus collected value of shipments data will help us evaluate the effectiveness of using the former as a proxy for the latter.

#### B. Quality Assurance

A quality assurance project team has been created recently, with the goal of implementing a QA system on all aspects of production.

#### C. Management Information System

Within the project, an effort is being made to construct a minimal management information system to track cost. If that information could be related to benefits derived from a specific activity within the project, then a great deal will have been accomplished toward making more informed decisions.

#### D. Resampling

Once Mining and Manufacturing has been completely surveyed, a recycling process will be implemented. We plan to resurvey each industry every five years. This is expected to keep costs manageable and the estimates reasonably in line with reality.

At this time several alternatives suggest themselves for this resampling. "Keyfitting" can be considered for the purpose of reducing initiation cost by increasing the chance of retaining current sample members. Rotational sampling (creation of panels) could be a new design if the respondent burden is considered to be a serious problem. The simplest solution operationally would be to select new independent samples, disregarding the fact that a unit was in the previous sample for the industry.

#### E. Composite Estimation

Composite estimators have been used successfully in several statistical programs in recent years (e.g., the Current Population Survey and the Consumer Price Index Revision). Perhaps it would be an applicable and beneficial technique in this program. We need to investigate this possibility.

#### REFERENCES

- [ 1 ] Early, J.F., Improving the Measurement of Producer Price Change, Monthly Labor Review, April 1978, pp. 7-15.
- [ 2 ] Early, J.F., Producer Price Indexes: A New Analytic Framework to Replace the Wholesale Price Index, American Statistical Association, Business and Economic Statistics Section, Proceedings, 1978.
- [ 3 ] Tibbetts, T.R., An Industrial Price Measurement Structure: The Universe Matrix of Producers and Products, American Statistical Association, Section on Survey Research Methods, Proceedings, 1978.

#### DISCUSSION

Chester H. McCall, Jr., M/M Associates

I appreciate the efforts made in the papers to develop a set of understandable definitions for the universe of interest, basic analysis units (the production unit), and the sampling frame. Conceptually, I found the papers sound, logically developed, and easy to comprehend. Perhaps the most appalling point (buried somewhat in both papers) was the admission that today's Consumer Price Index, published by the Bureau of Labor Statistics and a measure with which many labor contracts rely for cost of living increases, is not based upon sound statistical principles but rather depends heavily upon commodity experts in the Bureau. It is gratifying to learn that efforts are being made in developing the new indexes to correct this serious flaw.

There are too many initials used in the papers without proper definitions. The non-Bureau of Labor Statistics person would have trouble with the PPIR, WPI, ISPI, POPI, PPI, PFU, and CPIR initials.

And, finally, the papers generate quite a few questions which must be answered before the revised indexes can be considered as statistically based: what about the variance estimates? what was specifically learned about estimation problems? what are the criteria for sample size determination?

Considering the significance of the current indexes, let's hope that the revisions will have some sensible measures of statistical reliability.