AN INDUSTRIAL PRICE MEASUREMENT STRUCTURE: THE UNIVERSE MATRIX OF PRODUCERS AND PRODUCTS

T.R. TIBBETTS, BUREAU OF LABOR STATISTICS

SURVEY OBJECTIVES

A new industrial price survey has recently been designed at the Bureau of Labor Statistics to provide improved estimates of price change. The improvements include (1) a firmer theoretical foundation for the indexes, (2) the introduction of probability sample techniques, and (3) greater analytic usefulness and compatibility with other economic data. This paper is primarily concerned with the impact of these objectives on sampling.

Index Applications. Price indexes serve several functions and each must be considered in a survey design. Some of the functions, viewed separately, may suggest differing universes and estimating structures. Periodic surveys of data users 1/, as well as specific user requests have shown the following types of applications of existing industrial price indexes:

1. Measurement of general inflation. A broad indicator is needed to gauge aggregate price behavior throughout the economy.

2. Informing economic policy decisions. Government needs price measures for major economic sectors in order to form tax policies, interest rates, budget levels, and tarrifs.

3. Deflation of current dollar values to real terms. Estimation of real output in the national accounts requires indexes of price change,classified consistently with the accounting structure.

4. Forecasting of future price trends -- both at the manufacturing level and to anticipate passthrough to consumers.

5. Market analysis in the business community. Detailed estimates of relative price change for products and commodities are used to plan marketing strategy and to evaluate a firm's performance against industry-wide pricing.

6. Price adjustment (escalation) clauses in purchase and sales contracts. Over $100 billion 2/ is currently subject to adjustment by BLS indexes.

Different Applications Assume Different Universes. There are other applications, but these illustrate the range that a survey designer must consider in developing an estimating model. For the purposes of this discussion,the applications may be treated in two classes: those which require a producing industry orientation and those which require commodity or product orientation regardless of the classification of producing industries. This distinction is illustrated by the question: Are the most useful estimates those for the entire output of the dishwasher industry which is comprised of establishments producing dishwashers and several other products, such as clothes washers, or should the estimates reflect all of dishwasher production from establishments in any industry? The former case would estimate a universe of producers; the latter would estimate a universe of products.

Industry Universe of Producers. An industry orientation is appropriate if the estimate is intended for macro-economic analysis which involves, for example, wage and productivity data. These data relate to total establishment activity, and would generally be analyzed in an industry context. Econometric models for forecasting or identifying components of aggregate value require price change variables for the same establishment universe as the wage, productivity, capital, and other variables. Similarly, estimation of components of gross domestic product in the national accounting system requires that the price deflators match the industry classification of the accounts. A price index for dishwashers made in any industry is not the most appropriate deflator for the dollar value of the dishwasher industry.

One of the most useful tools for analysis of price behavior is the Input-Output (I/O) model 3/ of economic activity. Various stages of processing -- from raw materials to finished goods -- are identified and indexes of price change are estimated for each stage. Then, the pass-throughs and ripple effects of price changes can be anticipated and traced. The Input-Output framework reflects flows among industries and requires an industry classification, rather than a product orientation. Thus, the price index structure needed to support stage-of-processing analysis of price behavior must be based on or be matchable to the I/O industry structure.

Another level of application is to determine overall, aggregate price trends in the economy. This requires highly aggregated price measures for sectors such as Total Manufacturing, Total Durables, or Total Energy. These totals must avoid duplication or multiple weighting of the component price changes 4/. For example, a Total Energy index should reflect an increase in the price of crude oil only once; not successively as that increase is passed through to the refined product and to electric power. The increases for subsequent stages should be counted in the Total Energy index, i.e. weighted, only to the extent that the increased cost of the input has been netted-out. The appropriate weights for components of an aggregate index are net output values. Again, the price index structure, including weights, that can support this net output requirement of aggregate indexes is an industry structure. The flows which determine net output weights are among producers and industries, not among groups of products.

These illustrations describe the basis for a producing industry orientation of a price measurement design. But the objectives which lead to an industry orientation are only part of the applications previously noted. Other applications lead to a different orientation.

Product Universe. For some purposes, price indexes based on sampling of products and commodities made in any industry -- and for groupings of those products -- are appropriate. Increasingly, delivery of goods is contracted for far in advance of delivery, especially when reliability of supply is essential for the buyer (as for fuel supplies to a utility company), or when manufacturing takes a long time (weapons procurement). Especially during inflationary periods, these contracts include price adjustment clauses, called escalators, which specify that the price to be paid will be determined, in part, by the movement of a particular BLS price index. Here, the measurement system has a large and direct impact on the phenomenon being measured. In order to escalate materials costs, wherever-made product indexes, regardless of producing industry, are the best components of a composite escalator.

Other product-oriented applications are market analysis, assessment of purchasing efficiency,cost forecasting, and inflation monitoring and control. These applications, as well as escalation, require detailed product indexes which are not necessarily intrinsic outputs of an industry survey. The level of publishable product detail needed to support these applications cannot be developed efficiently by simple sampling among producers' total outputs.

The product-oriented applications are useful primarily to business, particularly to manufacturers. These firms also are the suppliers of the data. The success of a voluntary survey, especially one which elicits sensitive, proprietary information, is highly dependent on the extent to which the estimates meet the needs of respondents. Therefore, any preference of macro analysts, particularly in academia and government, for an industry orientation must be balanced against the preference of potential respondents whose cooperation is essential for any survey design to succeed.

SURVEY DESIGN REQUIREMENTS

To meet the dual objectives of producer estimates and product estimates, the designers of the new survey have examined differing universe definitions and sample frame possibilities. In the purely abstract sense, the universe of interest to price measurement is the universe of transactions in which prices are formed. In order to be measured, the transaction universe must be related to a data source -- a party to the transaction -- and must be structured into measurable units such as products, units of production, purchase orders, and shipments.

Production Units To Be Represented. From the industry perspective, the conceptual universe is composed of production units. Theoretically, a production unit consists of plant and equipment, technology, labor, and entrepreneurship which can be defined and specified -- that is, the value of each production variable is known and is under the control of management. Changes in the values of the variables produce predictable changes in cost and output. This production unit can be viewed as a "firm" in the abstract model 5/ developed in the "Theory of the Firm." Production units are intuitively attractive as a conceptual universe because all transactions flow from them and can be associated with a particular unit, and because the theoretical production unit has economic properties which are consistent with the methodological constraints of price index construction.

The production unit is also attractive to the sampling statistician for a number of reasons.

1. It can be represented by real units, such as companies, corporate divisions, plants, establishments, or organizational subdivisions of these.

2. It has measurable characteristics, some of which may provide stratification variables, such as size, location, age, affiliation with related units, etc. Other essential characteristics are available for index estimation, such as type of output, quantity or value of output, and prices received.

3. Files exist for some of the theoretical unit's real world approximations, and those files contain some of the characteristic information mentioned above.

Producer Frame Sources. Gaining access to frame sources to represent a universe of production units is quite difficult. The Census Bureau's quinquennial industrial censuses 6/ create a file of all producing establishments in the U.S., with detailed information about quantities and values of specific outputs, total shipments values, names, addresses, etc. The published Census reports are compiled from these files,and portions are updated annually through the Survey of Manufacturers. This file would be an ideal raw source for constructing production unit frames. Unfortunately the Census files are statutorily restricted to use by Census employees for compiling summary reports. Since this file is unavailable to the BLS, other frame sources have been used.

Investigation has shown that the best available source for a universe file of production units is the Unemployment Insurance (UI) file which is compiled by each of the 50 states and consolidated at the BLS 7/. If perfectly maintained, this file would be a complete list of all producing establishments in the U.S., with establishment name, address, employment size, and industry classification. The weaknesses of the UI file for price sampling are the absences of major product identification, of an output value measure, of parent company names, and some inconsistency among the states in quality and reference period.

Several commercial files of companies and establishments are also available. Some of these files include information for each unit on major product lines, secondary products, values of outputs, and parent company hierarchies. These advantages are offset by some incompleteness, misclassifications by industry,and substantial variation in reference period. The high cost of a commercial file also is a disadvantage. For the producer price survey, this type of commercial file is used to refine the UI file,along with information developed in direct contact with industry associations and major producers. The refinement objectives are to improve file accuracy, and to assemble better approximations of production units than appear in a simple list of establishments. Production units may consist of clusters of establishments belonging to the same company, when such clusters are managed as a unit 8/.

Development of industry-oriented universe files of production units is, therefore, complex but manageable. The parallel objective of the design -- to develop a frame for efficient estimation of detailed product indexes -- is more difficult. If product indexes were the only objective, the sample frame would reflect a universe of commodities and products. The producer universe also must be incorporated when the objective is to represent both producers and products. Theoretically, the universes of producers and of products should be consistent and complementary.

Products To Be Represented. The universe of products is theoretically comprised of the myriad individual physical outputs of mining and manufacturing. These outputs must be defined and classified so that realistic sampling units can be identified. That is, "product" must be defined to distinguish a particular unit from among the spectrum of item levels. These levels range from homogeneous classes of products to very detailed specifications of a product model and its terms of sale. For example, product may be defined at a broad level such as "autos" or it may be limited to a particular specification such as a 1978 National Motors Panther Model Z6, with options 1, 2, 5, and 7, for sales of that model in California to Rent-A-Car. Is the appropriate sample unit "auto" or "Panther" or a particular Panther prepared for the California fleet market? How detailed a product specification is appropriate to describe a primary sample unit?

Product Frame Sources. In addition to the conceptual formulation of a universe of products, consideration must be given to (1) available data sources, and (2) index users' demand for product detail. The former constraint often leads to a fairly broad definition of product, while the latter demand calls for greater detail. Available frame sources are limited. They include the published census product detail, which are often classes of items somewhat like the conceptual unit but which vary greatly in level of detail from industry to industry. The Census is the most complete and consistent source, however. Greater detail can be developed for fragments of the universe from unpublished Census detail, Commodity Classification lists, product directories for particular industries and commercial catalogues. Rough frames can be constructed for products, but they lack rigorous completeness, consistency, and value data. A lack of value data is serious since a basis for product sampling proportionate to size is needed.

HISTORIC APPROACH

These weaknesses have plagued the historic industrial price index program, which has been product oriented. The Wholesale Price Index (now titled Producer Price Index) is a venerable survey which was started at the beginning of the century and has been expanded over the years so that it now includes about 2800 commodity indexes based on about 10,000 price quotations 9/. From its inception, the WPI has been product oriented: the universe is defined as all commercial transactions of products produced by or imported into the domestic economy; the primary sample unit is a product specification; the aggregate indexes are product groupings, and the total is titled "All Commodities;" and the WPI weights are total shipment values of products, wherever-made. The majority of users of the WPI are business people, especially manufacturers, whose applications of the indexes are product-oriented 10/.

The sample of WPI products has been selected judgmentally by BLS commodity experts. The main criteria are to represent products with the largest shipment values and to price those products from major producers. The resulting indexes are dominated by volume-selling products of large companies. In recent years, some representation of smaller companies has been added to the WPI.

This traditional approach has produced a very useful measure of general price trends for important groupings of products, and has provided a large number of detailed product indexes for market analysis, and contract escalation. The WPI has been widely used for deflation and in macro-econometric models, with predictable difficulty in these industry-oriented applications. The absence of scientific sampling has precluded the calculation of measures of precision, reliability, variance, or error 11/.

In order to meet the need for industry-oriented indexes, the detailed WPI product components have been reclassified and reweighted along SIC lines to produce the Industry-sector Price Indexes (ISPI) 12/. This makeshift program was initiated in 1965 with indexes calculated retroactively to 1957. Lack of funds restricted this effort to use of the available WPI data base. The ISPI's price only about one-fourth of the total value of the mining and manufacturing universe. Reweighting was limited by the level of published information in the industrial censuses.

The ISPI's represented a modest step toward inclusion of industry and producer estimates in the pricing program. However, these indexes suffer from the same lack of sampling technique as the WPI. Furthermore, the particular prices borrowed from the WPI were not originally selected to represent a producers' universe.

THE REVISED DESIGN

The need for a new industrial price measurement approach is clear. It is also clear that a new design should address the needs of both producing-industry-oriented users of the data and those who need detailed product indexes. Those respective universes have been described. The first step in designing a new approach was to specify an economic concept for integrating these universes and providing an estimating framework. At the same time, the concept should lead to a definition of "true" price change for a fixed unit of production and for a constant product, in contrast to price changes caused by changes in the product or in production technology.

Economic Concepts. The most useful concept is that of the production function, combined with the "Theory of the Firm" 13/. The theoretical firm's goal is profit maximization. In the most oversimplified terms, Profit = Revenue - Costs. Revenue is the product of the quantity and price of output. Cost is the sum of expenditures for material and labor inputs, plus depreciation of plant and equipment. In the model, perfect competition is assumed and price is determined externally by the market. For particular production units, it is reasonable to expect that profit maximizing decisions are fairly restricted and short-run, and cannot affect basic types and quantities of production inputs. Therefore, the price index concept assumes that types and quantities of inputs are fixed at the levels which existed in the index reference period. If plant, equipment, technology, and input are held constant, profit may be increased only by adjusting output in response to market conditions so long as unit revenue exceeds unit cost. The price changes which result from such profit maximizing decisions are the relevant changes for index measurement. The theoretical output price index represents the percentage change in revenue received by the production unit as the result of a change in prices. Price changes which result from improved technology, changes in input, different quality of the product, or market shifts to different producers should be excluded from the index in order to isolate "true" price change, as received by particular producers. The underlying concept helps define the exclusions to "true" price change 14/.

The Universe Matrix Model. Implementation of this general price measurement approach requires, first of all, the specification of a universe with the dual characteristics described above. The ideal sample frame includes identification and output value for all producers classified by producing industry, and for all products wherever-made. Although it is not currently possible to construct such a matrix in detail because of its immense size and because much of the data have not been collected, it is nevertheless useful to specify the model. The theoretical matrix provides a framework and planning tool for designing real frames and samples.

The matrix model is simple and obvious; the effort to construct the actual matrix is more interesting and useful. (See Fig. 1) All production units are listed as the stub and grouped by producing industry classification (SIC). All products are listed horizontally as column headings, grouped by primary producing industry. The very small economy depicted in Figure 1 is composed of 9 producers (A-I), grouped in three industries (I-III). The output of this economy is classified into nine

MODEL UNIVERSE OF PRODUCERS AND PRODUCTS

OUTPUT VALUE, CLASSIFIED BY INDUSTRY
(thousands of dollars)

Industry → / ↓ Producer	I		II (PRODUCTS)				III			Total Producer Output	Total Industry Output
	1	2	3	4	5	6	7	8	9		
I	300	100									515
A	200	75		(25)	(50)					350	
B	100						(20)			120	
C		25	(20)							45	
II			40	30	70	20					220
D				10	20	10		(10)		50	
E		(40)	30	20						90	
F					50				(10)	60	
G			10			10				20	
III							1200	200	100		1700
H				(100)		(100)	900			1100	
I							300	200	100	600	
Total product output, wherever-made	300	140	60	155	120	120	1220	210	110		2435

FIGURE 1

"products." These products are classes of items as described previously. Completing the matrix would require entering estimates for shipment values in the appropriate cells. Not all cells are to be filled since producers specialize and industries are defined according to those specializations. At a minimum, each producer (row) has one product output and each product (column) has at least one producer. Most often, a producer makes several products and each product is made by several producers. Thus, the value entries begin to form sets grouped around a diagonal line.

If producers made only products which are primary to their industry, as specified in the industry definition, the diagonal pattern would hold and total output for an industry would equal the total value of products of that industry. Then the survey designer could partition the universe into industries and adequately sample both producers and products from the same industry frame. However, producers don't behave to suit survey designers; they perversely produce products "belonging" to other industries. For example, producer A may be primarily a chemical company which derives three-fourths of its revenues from chemical products 1 and 2. Thus, A is correctly classified in the chemical industry. However, A also produces some fertilizers -- products 4 and 5. The value of A's fertilizers are part of the total output of the chemical industry, and are part of the total value of fertilizer output, wherever-made. This type of cross-over defines the important difference between industry output indexes and product indexes -- both for sample design and for weights and prices used in the index estimates. A straightforward sampling of units in either dimension would not automatically and efficiently produce good estimates for units in the other dimension.

Use of the matrix approach helps clarify the issue; it allows for approximating the magnitude of the cross-over problem for particular industries and for identifying difficult product areas. The conventional term for production of products outside the producer's industry classification is "secondary production." The percentage of total industry output included in primary production, that is, of products classified as belonging to the industry, is termed the "specialization ratio."' Value estimates of secondary production and specialization are available from the industrial censuses for industry totals at the 4-digit SIC level 15/. In Figure 1, these values can be derived by comparing the underlined figures (primary) with the figures in parentheses (secondary).

The two dimensions of the matrix (Fig. 1) reflect the classes of survey objectives described above:

1. Industry/producer estimates to support macro-analyses or industry-oriented applications such as deflation of industry values, econometric combination with other industry data, or inflation monitoring of industries.

2. Product/commodity estimates to support micro-analyses or product-oriented applications such as materials escalation, market analysis, or forecasting and comparison of purchase costs.

Problems in Matrix Construction. Some problems with a comprehensive and detailed construction of the matrix should be noted. The size of the completed matrix would be very large. Instead of the nine producers in three industries shown in the model, the real economy has about 350,000 producers in 500 mining and manufacturing industries according to the Standard Industrial Classification 16/. Instead of the nine "model" products, the Census of Manufactures lists about 11,000 composite products. Even if all of the data were easily available, it is doubtful that the modest resources available for the survey would support construction and maintenance of an 11,000 x 350,000 matrix.

More certainly, available resources would not support collection of the missing data. As noted previously, some information is available. High quality data for 4-digit industries (Rows I-III) are available from the Censuses. Detailed data are also available for many (but not all) products, including primary and secondary, but only at the 4-digit level. The producer detail is not releasable by the Census Bureau, although much of these data exist on Census computer tapes. For the producer detail, the BLS has had to rely on other sources, which introduces the usual problems of matching, missing data, classification and definitional inconsistencies, etc..

The combination of various sources provides preliminary sample frames which list producing establishments, their industry classification, corporate relationships, and some information on major product lines. The total value of their respective outputs is partially available or can be represented by employment as a proxy. The matrix can be filled in for industry and product totals, and for estimated producer values for total output. Additional information from industry associations, special studies, and corporate annual reports can provide guidance for the "producer-by-product" cells, but these data do not support rigorous completion of the matrix. This deficiency affects the sample design strategy 17/.

Qualitative Requirements. A survey design that can meet both types of objectives must satisfy the qualitative requirements of estimates in both dimensions of the matrix. To specify these requirements, it is useful to return to the original objectives of the survey and assess the relative importance of differences in quality or completeness of data in different portions of the matrix. Since the survey's primary objective is to create Producer Output Price Indexes, one requirement is that the sample frame should accurately reflect the universe of production units. The frame should include accurate and complete lists of producers' names and addresses, accurate industry classification, and appropriate measures of sizes. These requirements are essential to support the primary estimates desired -- indexes of prices of producers' output at the industry level.

It is desirable that the quality of estimates for products also be high, but the constraints are less rigid. The impact of undersampling for specific products is limited to the detailed product estimates and need not affect the quality of the industry estimates. Conversely, selective enhancement of particular product samples will not improve the industry estimates greatly. Therefore the quality of sampling for specific products can be permitted to vary according to the degree of user interest in the product estimates and users' cooperation in providing detailed product and value information.

The specific objective, with respect to products, is to achieve good (i.e. low variance) estimates for the approximately 6,000 products with largest value, and to expect, initially, poorer estimates for the less important products. By sampling products proportionate to output value, and by controlling the sample allocation, in part on the number of products per industry, the design should generate "good" estimates for the large value products. Subsequently, samples for poorly estimated products can be supplemented as needed or as resources permit. Since the supplements have not been part of the sample design allocation, they will not improve industry estimates. Supplementation will simply allow publication of more reliable product indexes, which is the objective in the product area.

CONCLUSION

This paper has described the combined universe of producers and products. Some problems with estimating price changes from the universe have been discussed and frame sources have been outlined. Survey objectives have been related to user-oriented application of price indexes. Potentially conflicting elements of survey design have been noted, with indications of the priorities of objectives. Based on these requirements, the BLS has designed and tested a revised producer price index program (PPIR). For a description of the survey design and a discussion of statistical problems encountered, see "Survey Methods and Theory of the Industrial Price Revision" by Kulpinski, Cohen, and Perez-Lopez 18/.

The pilot survey, which formed the major test of the PPIR design, was completed in early 1978. Four industries were selected for a prototype survey which involved all major functions of the sampling and estimation design. Publication of the test indexes began on a monthly basis with release on August 10, 1978 of data from December 1977 through July 1978. Analysis of the results of the pilot survey is beyond the scope of this paper. In general, the survey met all of the fundamental objectives previously stated. The universe of producers seems, upon inspection, to be well represented, and samples for most major products are large enough to warrant publication. Variances have not yet been calculated.

The full scale survey, modified from the pilot experience, was started in June. Publication of indexes will begin in mid-1979 for the first sample of seven industries. Thereafter, at about three-month intervals, progressively larger groups of industries will be surveyed until all of the Mining and Manufacturing sector has been covered. If BLS budget requests are fulfilled, the PPIR should be completely initiated within five years.

FOOTNOTES

1. Dereck Moeller, "The BLS Industrial Price Program: A Survey of Users," Report 509 (Bureau of Labor Statistics, 1977).

2. *Ibid.*

3. "Input-Output Structure of the U.S. Economy: 1967," Supplement to the Survey of Current Business, 1974. The major adjustment needed to integrate I/O and SIC structures is to account for I/O's transfer of secondary production values to the products' primary industry.

4. William Nordhaus and John Shoven, "Inflation 1973: The Year of Infamy," Challenge, May/June 1974, pp. 14-22.

5. Franklin Fisher and Karl Shell, The Economic Theory of Price Indexes, New York, 1972.

6. Census of Manufacturers, 1972, (U.S. Bureau of the Census, 1976).

7. "Chapter 9. Employment and Wages Covered by Unemployment Insurance Laws," BLS Handbook of Methods, Bulletin 1910, (Bureau of Labor Statistics, 1976), pp. 66-73. Employment and Wages, (Bureau of Labor Statistics, quarterly).

8. Robert B. Archibald, "On the Theory of Industrial Price Measurement: Output Price Indexes," Annals of Economic and Social Measurement, Winter 1977, pp. 57-72.

9. "Chapter 14. Wholesale Prices," BLS Handbook of Methods, Bulletin 1910, (Bureau of Labor Statistics, 1976), pp. 109-22.

10. Moeller, *op. cit.*

11. Richard Ruggles, The Wholesale Price Index: Review and Evaluation, (Council on Wage and Price Stability, June 1977).

12. "Chapter 15.Industry Price Indexes," BLS Handbook of Methods, Bulletin 1910, (Bureau of Labor Statistics, 1976), pp. 123-26.

13. Fisher and Shell, *op. cit.*

14. John F. Early, "Improving the Measurement of Producer Price Change," Monthly Labor Review, April 1978, pp. 7-15.

15. Census of Manufacturers, 1972, *op. cit.*

16. Standard Industrial Classification Manual,1972, (Office of Management and Budget, 1972).

17. S.K. Kulpinski, S.J. Cohen, K.G. Perez-Lopez, "Survey Methods and Theory of the Industrial Price Revision," paper presented at meetings of American Statistical Association,San Diego, August 16, 1978.

18. *Ibid.*