

## **B. Collection of Information Employing Statistical Methods**

### **1. Universe and Sample Size**

The Producer Price Index is an on-going survey of net transaction prices received by producers. The current universe for the PPI survey consists of roughly 5.0 million establishments comprising the covered portions of the mining, manufacturing, forestry, utility, and non-goods producing sectors. About 5,836 respondents are initiated in any given year, while 26,250 respondents provide monthly data pertaining to 105,000 price quotations. The replenishment and rotation of respondents within the PPI occurs at a rate of approximately 13.3 percent each year. The list containing the universe of all producing establishments comes mostly from Unemployment Insurance (UI) files provided by state agencies. Supporting information and alternative frames may be obtained from other sources, if they are deemed to be more accurate. Total initial cooperation is about 80 percent: Roughly 20 percent of establishments slated for initiation into the PPI refuse to cooperate. Frame error further reduces the PPI's initial response rate, while respondent attrition over the life of the sample affects the long-term cooperation rate.

### **2. Collection Procedures**

The PPI survey is based on probability-proportional-to-size sampling. Every establishment listed as belonging to the universe of entities producing in the to-be-sampled NAICS industry, regardless of size, has a chance of being selected. The chance of any single establishment being chosen for participation in the survey is commensurate with its importance to the industry as a whole. Comprehensive coverage is necessary to insure that the price data collected is a representative sample of the universe of pricing activity within an industry. It is the PPI's opinion that the burden imposed on business establishments is very near the practical minimum consistent with production of a statistically meaningful index.

The steps involved in probability-proportional-to-size sampling include: constructing a frame (a list of businesses from which a sample is to be selected), identifying any specific variables that represent unique price-forming groups (explicit stratification), calculating the number of sample units and price quotations required within each unique group, sorting each group by a measure of size (usually employment), and using a calculated sample interval to select a representative subset of entities from the list. Probability-proportional-to-size sampling, in addition to improving efficiency and reducing bias, provides the capability to calculate statistical estimates of reliability, precision, and error.

The number of establishments and price quotations selected for tracking varies, depending on the homogeneity within the sampled industry. The sample must be large enough to represent the full range of producers and products. Since participation in the survey is voluntary, not every entity selected for inclusion cooperates. Furthermore, sample frames typically contain a certain degree of error. Frame error includes entities

which are out of business or those incorrectly classified. Anticipated respondent attrition over the life of the sample also influences sample allocation.

Once a respondent has been approached by a BLS data collector and agrees to cooperate, initiation into the PPI survey requires, on average, 2 hours of respondent time. The amount of time varies, however, depending on the manner in which company records are kept, the number of price quotations requested, and the nature of the industry and products being sampled. The first step in initiating an establishment into the PPI involves verification of address and employment information. The next step involves identifying product lines produced or service lines provided, along with revenue data for each activity. The third step is item selection, which BLS refers to as disaggregation.

For each line of activity, respondents identify unique price-determining characteristics that come into play, along with the revenue that each line generates. A random number table is used to choose the unique transactions that will be tracked by PPI. This process is repeated for increasingly detailed categories until completely unique transaction types are identified. Disaggregation identifies unique price-determining variables, both product and transaction specific, and assigns a weighted importance to each. Identifying unique activities and their importance relative to the respondent's full revenue-generating activity allows the PPI to efficiently sample a representative subset of transactions, and permits efficient recording of these classification parameters for future tracking. The BLS National Office provides forms to data collectors to assist in the process of assigning probabilities, selecting transactions, and documenting sampled transactions. (See attachments: forms BLS-1810A, BLS-18A1, BLS-1810-B, BLS-1810C, BLS-1810-C1, and BLS-1810E.)

Effective with the release of data for January 2004, the PPI converted its sampling, data collection, and industry-based publication structures to the NAICS. Through December 2003 PPI's industry-based procedures were linked to the SIC organizational system.

During monthly repricing, the main communication tools between the PPI and respondents are the price-collection form BLS 473P (See attachment) and the BLS Internet Data Collection Facility (IDCF). The Program currently sends out approximately 105,000 pricing requests per month to roughly 26,250 responding establishments. One request exists for each price quotation that is being monitored. The request contains the specific information required by the PPI to track changes in net transaction prices for predetermined outputs. Survey requests are designed to take industry-specific factors into account, allowing adaptation to individual company accounting and data structures. The program has modified the form, for implementation in March 2014, in order to streamline and simplify the layout and content. The changes in the form are expected to result in not only more accurate data but also a significant decrease in burden. As mentioned in section A.3, Technological Collection Techniques, the Program has introduced a new capability for survey participants to provide monthly repricing updates over the Internet. The online screens are structured similarly to the price-collection form BLS 473P. (See attachment for screen shots.) As usage of the

Internet facility grows, the number of forms sent out via the mail or fax will decline, though the number of requests for updated information will remain the same overall.

When price-quotation questionnaires are returned by fax or mail, they are entered into a database using an optical scanner. Respondents often submit forms that include changes to product descriptors, transaction descriptors, or net transaction prices. These changes may require a telephone call from a PPI industry analyst for clarification and verification. With monthly repricing via the internet, data reported by respondents is automatically transferred from the BLS Internet Data Collection Facility to the PPI database on a daily basis with items requiring follow-up by BLS staff flagged by our computing system.

Detailed-level price indexes are constructed by combining price quotations from respondents that describe similar product or service categories. Aggregate indexes -- whether they are product line, industry, industry group, commodity group, or final demand-intermediate demand -- are weighted averages of detailed-level price indexes.

The modified Laspeyres formula below approximates the actual computation procedure for Producer Price Index:

$$I_t = \left[ \frac{(\sum Q_a P_0 (P_t/P_0))}{(\sum Q_a P_0 (P_{t-1}/P_0))} \right] \times I_{t-1}$$

where  $P_0$  is the price of a product in the comparison period,  $P_t$  is the current price, and  $Q_a$  represents the quantity shipped during the weight-base period. In this form, an index is the weighted average of price ratios for each item ( $P_t/P_0$ ) in a detailed cell.

Within each PPI detailed cell, individual price quotation reports from establishments are given different weights, in accordance with data on shipment values provided to BLS field representatives by respondents during initiation interviews, adjusted by BLS using probability selection techniques.

If a price quotation report has not been received in a particular month, then the change for that price will in general be estimated by averaging the price changes for the other items within the same detailed cell (that is, for the same kind of products) for which price reports have been received.

### 3. Methods to Maximize Response Rates

Four months after first publishing its set of monthly indexes, PPI recalculates and finalizes indexes, taking into account late reports and back-corrections received from respondents. At this four-month mark, approximately 70% of price-quotation questionnaires are returned.

In order to maintain and improve cooperation, the PPI maintains a procedure that includes contacting, by telephone, any selected respondents that have not returned forms for a specified period of time. Assistance is provided with regard to any aspects of the form that at first glance appear unclear or burdensome; a common reason for non-response. Note that since PPI started expanding into non-goods producing industries, the program has experienced generally lower response rate for non-goods -producing industries than for goods-producing industries, impacting overall PPI response rates. Non-response can occur due to refusal to cooperate or due to frame error, and both situations are more common with non-goods producers. In the past, the number of attempted annual initiations increased to compensate for a lower expected response rate.

In March 2012, BLS shared PPI's final report with the Office of Management and Budget on an analysis to determine if non-response bias exists in its published data. The study consisted of three phases. Phase I attempted to determine if non-response in PPI data is systematic or random. The focus was on the two primary stages of data collection -- initiation of primary sample units and monthly repricing of reported data. The modeling of unit initiation sought to identify correlations between unit response at initiation and certain frame variables. The modeling of (unit and item) repricing response sought to identify correlations between unit/item repricing and certain frame variables. Phase II of the study focused on the relationship between any item short-term relatives (STR's) and the frame variables found to be significant in phase I. We sought to determine if prices trend differently for the variables identified as being correlated to response. Phase III sought to determine if non-response bias has a statistically significant effect on PPI indexes. The Report contained the following conclusion:

*The objective of our research was to determine if nonresponse bias existed in PPI data. We found a statistically significant correlation between response and several variables including Employment, Region, and Shipments and Receipts. Using Shipments and Receipts, and not Employment, as a measure of size, we found a statistically significant association between item prices and Region and Shipments and Receipts. We excluded Region from further consideration (due to its very strong correlation with industry). We used Shipments and Receipts to create nonresponse adjustment cells and calculated item weights that account for nonresponse. We used these adjusted item weights to calculate adjusted indexes. Original and adjusted index estimates were calculated and compared to see if there were statistically significant differences between them. Our analysis showed that very few PPI indexes exhibited signs of nonresponse bias and the very few that did were affected by unusually high nonresponse in very specific size classes. These findings did not identify strong evidence of nonresponse bias in PPI indexes for the industries and years that were analyzed. There is no need to adjust for nonresponse on a systematic basis. Even so, PPI will continue to stress its ongoing efforts to improve response such as monitoring response more closely and emphasizing repricing delinquency follow-up. With continued care in the selection of industries to be resampled, improvements to our sampling methodology, and concerted nonresponse follow-up efforts, nonresponse can be kept at an acceptable minimum.*

#### 4. Testing Procedures or Plans

The PPI is planning to test some new questions in regard to the implementation of the redefinition of manufacturing under NAICS in the next few years. Currently, the plan for the redefinition of manufacturing, including the identification of factoryless goods producers, is to begin implementation in affected Federal statistical programs with the publication of NAICS 2017. A non-substantive change further detailing PPI's testing plans will be submitted to OMB for approval prior to any testing.

#### 5. Statistical Contacts

Oversight of statistical methods in the PPI survey are maintained by the Bureau of Labor Statistics, Office of Prices and Living Conditions, Division of Price Statistical Methods, Steven P. Paben, Supervisory Mathematical Statistician, (202) 691-6147.

### C. PPI Methodology References

The methodology of the PPI has been documented in numerous papers and articles written since 1977 when the PPI underwent the most comprehensive redesign in its history. These papers cover a broad spectrum of topics ranging from price theory and program concepts to actual data collection methodology. A list of references includes:

Archibald, Robert B. "On the Theory of Industrial Price Measurement: Output Price Indexes," *Annals of Economic and Social Measurement*, Winter 1977.

Bureau of Labor Statistics, *BLS Handbook of Methods*, U.S. Department of Labor. Available at <http://www.bls.gov/opub/hom/homtoc.htm> Chapter 14

Buszuwski, J.A. and Scott, S. (1988), "On the Use of Intervention Analysis in Seasonal Adjustment," *Proceedings of the Business and Economics Section*, American Statistical Association.

Buszuwski, J.A., (1987) "Alternative ARIMA forecasting horizons when seasonally adjusting producer price index data with X-11 ARIMA in concurrent mode" ASA Proceedings of the Business and Economic Statistics Section.

Buszuwski, J.A., (1986) "Alternative seasonal adjustment forecast horizons and methods for the Producer Price Index" ASA Proceedings of the Business and Economic Statistics Section.

Buszuwski, J.A., (1993) "Some issues in seasonal adjustment when modeling intervention" ASA Proceedings of the Business and Economic Statistics Section.

Chen, Helen and Sadler, A. (2010) "Comparison of Variance Estimation Methods Using PPI Data" ASA Proceedings of the Government Statistics Section.

Colia, Demetra. (1988) "Measuring sample variability in the producer price index," ASA Proceedings of the Section on Survey Research Methods.

Council on Wage and Price Stability, *The Wholesale Price Index*, June 1977.

Early, John F. "Improving the Measurement of Producer Price Change," *Monthly Labor Review*, April 1978.

Gerduk, Irwin. (1984) "Quality assurance elements in Producer Price Index data initiation," ASA Proceedings of the Section on Survey Research Methods pp 151-156.

Hellerstein, Judith. (1989) "The effects of sample size on variances of the Producer Price Index," ASA Proceedings of the Section on Survey Research Methods pp 170-175.

Hill, Kimberley Dailey. (1987) "Survey Design in the Producer Price Index," ASA Proceedings of the Section on Survey Research Methods pp 583-588.

Kulpinski, Stanley; Cohen Stuart J.; Perez-Lopez Kathleen, (1978) "Survey methods and theory of the Producer Price Index revision," ASA Proceedings of the Section on Survey Research Methods pp 517-521.

Popkin, Joel. "Integration of a System of Price and Quantity Statistics with Data on Related Variables," *Review of Income and Wealth*, March 1978 pp 25-39.

Sager, Scott D. "Effect of 1992 Weights on Producer Price Indexes," *Monthly Labor Review*, July 1996 pp 13-23.

Sinclair, James and Catron, Brian. "New Price Index for the Computer Industry," *Monthly Labor Review*, October 1990.

Slack, David and Hagemeyer, Kirk (2007) Survey Response Measurement Team Quarterly Report.

Tibbetts, Thomas R. "An Industrial Price Measurement Structure: The Universe Matrix of Producers and Products," *1978 Proceedings of the Section on Survey Research Methods*. American Statistical Association, Washington, DC, 1979 pp 511-516.

U.S. Department of Labor, Bureau of Labor Statistics, *Escalation and Producer Price Indexes: A Guide for Contracting Parties*, Report 807. Original produced September 1991; last updated July 2006: <http://www.bls.gov/ppi/ppiescalation.htm>.

The Bureau of Labor Statistics in the Monthly Labor Review has published additional articles on specific PPI topics. A list of those articles includes:

- Analyzing price movements within the Producer Price Index Final Demand-Intermediate Demand aggregation system (January 2014)
- Comparing new final-demand producer price indexes with other government price indexes (January 2014)
- Scheduled passenger air transportation in the Producer Price Index: improvements and trends (September 2013)
- Price transmission effects through three stages of food production (December 2012)
- The behavior of the Producer Price Index in a global economy (September 2012)
- A new, experimental system of indexes from the PPI program (February 2011)
- PPI and CPI seasonal adjustment: an update (July 2010)
- New wherever-provided services and construction indexes for PPI (August 2009)
- Producer price highlights, 2008 (July 2009)
- Producer price highlights, 2007 (July 2008)
- Price measures of new vehicles: a comparison (July 2008)
- Producer price highlights, 2006 (July/August 2007)
- Price transmission: from crude petroleum to plastics products (December 2006)

- Price and expenditure measures of petroleum products: a comparison (December 2006)
- Producer price highlights, 2005 (July 2006)
- Price transmission within the PPI for intermediate goods (May 2005)
- Factors affecting beef and cattle producer prices movements (May 2005)
- Consumer gasoline prices: an empirical investigation (July 2003)
- An empirical analysis of price transmission by stage of processing (November 2002)
- Producer price highlights, 2001 (July 2002)
- Producer price highlights, 2000 (July 2001)
- Producer price highlights, 1999 (August 2000)
- Producer price highlights, 1998 (July 1999)
- Comparing PPI energy indexes to alternative data sources (December 1998)
- The 1996 grain price shock: how did it affect food inflation? (August 1998)
- Producer price highlights, 1997 (July 1998)
- Improving the PPI samples for prescription pharmaceuticals (October 1997)
- Producer price highlights, 1996 (July 1997)
- Producer price highlights, 1995 (July 1996)
- Effect of 1992 weights on Producer Price Indexes (July 1996)
- Hospital price inflation: what does the new PPI tell us? (July 1996)
- Producer price highlights, 1994 (July 1995)
- Pricing practices for tobacco products, 1980-94 (December 1994)
- Producer price rises slowed in improving economy in 1993 (May 1994)
- Effect of updated weights on Producer Price Indexes (March 1993)
- Recession and energy prices ease producer prices in 1991 (May 1992)
- New price index for the computer industry (October 1990)

PDFs of all the articles can be accessed from this web page: <http://www.bls.gov/ppi/ppimlr.htm> .

For 2010 and 2011, BLS published a series of quarterly articles on its price data including the Producer Price Index program in the series called “Focus on Prices and Spending”. The articles cover both topics of interest as well as price trends during the past quarter:

- The 2012 PPI Weight Update: How and why does BLS update revenue weights? (Fourth Quarter 2011)
- Technology in Commodity Markets (Third Quarter 2011)
- The Impact of Fuel Surcharges on the PPI (Second Quarter 2011)
- Grain Prices and Our Grocery Bill (First Quarter 2011)
- Producer Price Index Introduces Experimental Aggregation System (Fourth Quarter 2010)
- What Does the Producer Price Index Measure? (Third Quarter 2010)
- PPI Industry or Commodity Data: Which Better Suits Your Needs? (Second Quarter 2010)
- Intermediate Materials Continue their Advance; Other Producer Price Highlights (First Quarter 2010)

You can access PDF and HTML versions of the articles at: <http://www.bls.gov/opub/focus/> .



Beginning in May 2012, *Focus on Prices and Spending* has been consolidated with several other BLS publications into a new series of reports called *Beyond the Numbers*. *Beyond the Numbers* illuminates interesting, timely BLS data on discrete topics and is oriented towards a general audience.

- Improvements to the Producer Price Index measure: the Final Demand-Intermediate Demand system (February 2014)
- Highlights of the 2013 PPI User Survey (August 2013)
- The Effects of Shale Gas Production on Natural Gas Prices (May 2013)
- Gold prices during and after the recession (February 2013)
- Will the 2012 drought have a bigger impact on grocery prices than the 1998 drought? (November 2012)
- Wholesale And Retail Producer Price Indexes: Margin Prices (August 2012)
- How New Fees Are Affecting The Producer Price Index For Air Travel (May 2012)
- The 2012 PPI Weight Update: How And Why Does BLS Update Revenue Weights? (February 2012)

You can access PDF and HTML versions of these and other articles at <http://www.bls.gov/opub/btn/archive/home.htm>