

## 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. The PPI program initiates about 4,305 respondents annually. PPI repricing consists of 11,640 respondents providing data for 61,250 price quotations on a monthly basis. The PPI program began updating samples on a tri-annual basis with the release of the June 2015 indexes. Every sample update coincides with the release of data for February, June, and October. This approach to replacing industry samples results in representative indexes of current industry production. The PPI program constructs the list containing the universe of all producing establishments using Unemployment Insurance (UI) files taken from various state agencies. Supporting information and alternative frames may be obtained from other sources. The PPI utilizes these sources if they are deemed to be more accurate.

Note: the PPI does not distinguish between private and public establishments selected for the survey.

#### Initiation Reponses Rates

Initiation response rates for the PPI are computed based on unit (establishment) response at sample initiation.

Un-Weighted Initiation Annual (Fiscal Year) Response Rate =

$$\frac{\text{Number of establishments agreeing to participate in the PPI survey}}{\text{Number of establishments agreeing to participate} + \text{Number of establishments identified as out of business or out of scope during the initiation process}}$$

Note: establishments identified as out of business or out of scope during the initiation process are not counted in this response rate.

<b>Fiscal Year</b>	<b>Initiation Response Rate</b>
2019	73%
2018	75%
2017	76%

#### Repricing Response Rates

The PPI repricing response rate for estimation is an unweighted item based rate. The estimation response rate provides the percentage of items eligible for use in estimation and represents the actual response based on current potential for response. The numerator of this rate is the count of all items used in estimation and the denominator is the count of all items for which information was requested.

Un-Weighted Estimation Annual (Fiscal Year) Response Rate  $\hat{\rho}$

$$\frac{\hat{\rho} \text{ of items used } \in \text{ estimation}}{\hat{\rho} \text{ of items for which information was requested}}$$

Note: items that are out of season or have been discontinued are not counted in this response rate.

<b>Fiscal Year</b>	<b>Repricing Response Rate</b>
2019	76%
2018	74%
2017	79%

## 2. Collection Procedures

The PPI survey is based on probability-proportional-to-size sampling. Each establishment listed in the sampling frame of producing entities has a chance of selection for the PPI survey. The PPI program assigns this probability in proportion to the size of the establishment. The PPI program collects samples of populations of various individual NAICS industries. This means the chances of the PPI program selecting a given establishment for inclusion in the sample are proportional to the establishment's overall importance within its respective NAICS industry. 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 (stratum), sorting each group by a measure of size (usually employment), and using a calculated sample interval to

select a representative subset (sample) of entities from the list (frame). Probability-proportional-to-size sampling provides more reliable and precise estimates.

The number of establishments and price quotations selected for repricing 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, some selected establishments opt not to participate. Furthermore, sample frames typically contain a certain degree of error. Frame error includes entities defined as out-of-business, out-of-scope of the PPI, and those incorrectly classified. The PPI program anticipates respondent attrition over the lifespan of the sample. This expected attrition also influences sample allocation.

Once a selected establishment is identified and a respondent has been approached by the BLS, the data collector requests that the respondent agree to participate in the initiation interview. Prior to the onset of the coronavirus pandemic, PPI initiation interviews were primarily conducted in person by BLS Field Economists. Because of the coronavirus pandemic, Field Economists temporarily ceased all personal interviews and began conducting initiation interviews by telephone. PPI expects that other modes of collection, for instance, video will become available once security protocols are established. PPI anticipates that the next few sample updates will have fewer industries refreshed than normal, reflecting challenges related to initiating establishments into the survey, which normally involves in-person data collection. As a result, BLS may need to extend sample update cycles in order to collect enough data for some industries.

The average initiation interview requires two hours of the respondent's time. The first step in initiating an establishment into the PPI involves verification of address and employment information. The second 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.

Disaggregation identifies unique price-determining variables, both product and transaction specific, and assigns a weighted importance to each. 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 detailed categories until completely unique transaction types are identified. 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 forms: BLS-1810A, BLS-18A1, BLS-1810-B, BLS-1810C, BLS-1810-C1, and BLS-1810E.)

During monthly repricing, the main communication between the PPI and respondents is the BLS Internet Data Collection Facility (IDCF). PPI data collection using IDCF has replaced mail and fax as the main collection tool. Several months leading up to February 2018, PPI

staff worked with respondents to transition from mail and fax modes of collection to IDCF collection application. From February 2018 on, PPI has been a web-only survey. PPI arranged to obtain price data from mail/fax respondents who declined web repricing but were willing to provide data by other means. PPI now receives 99% of its respondent provided data through the IDCF application. Note that data acquired through a secondary source does not require respondent interaction and is not counted as part of the respondent burden information.

The PPI program currently sends out approximately 61,250 requests to about 11,640 respondents every month. The PPI program issues one request for each monitored price quotation. The request contains specific information for each price quotation being monitored. The PPI program designs survey requests to take industry-specific factors into account, allowing adaptation to individual company accounting and data structures.

With monthly repricing via IDCF, data reported by respondents are automatically transferred to the PPI database on a daily basis. Items requiring follow-up by BLS staff are flagged by the 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 provided below approximates the actual computation procedure for the Producer Price Index:

$$I_t = [(\sum Q_a P_o (P_t/P_o)) / (\sum Q_a P_o (P_{t-1}/P_o))] * I_{t-1}$$

where  $I_t$  is the price index in the current period,  $I_{t-1}$  is the price index in the previous period,  $P_o$  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 base period. In this form, an index is the weighted average of price ratios for each item ( $P_t/P_o$ ) in a detailed cell.

Within each PPI detailed cell, individual price quotation reports from establishments are given different weights, according to shipment values that respondents provide to BLS during initiation interviews. The weights are adjusted by BLS using selection probabilities.

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

### 3. Methods to Maximize Response Rates

Four months after first publishing monthly indexes, PPI recalculates and finalizes indexes, taking into account late reports and back-corrections received from respondents. The PPI

program receives over 70% returned price quotation questionnaires during this four-month period.

In order to maintain and improve cooperation, the PPI maintains a procedure that includes contacting, by telephone, any selected respondents that have not provided data for a specified period of time. Assistance is provided with regard to any aspects of data entry that at first glance appears unclear; a lack of clarity is a common reason for non-response. To further maximize response rates, PPI shares price information with other BLS price programs in certain situations. In an effort to increase efficiency and reduce overall respondent burden, the Consumer Price Index Program, the Producer Price Index Program, and the International Price Program may share resources to collect pricing information from respondents that are selected for inclusion in multiple surveys. In these cases, prices for the same product or service may be used by more than one price program; however, each program would determine appropriate weighting according to its own established methodology. All information shared across programs is used for statistical purposes only and is protected under the BLS confidentiality pledge.

The PPI conducted a study in 2012 to determine if non-response bias existed in its published data. Analysis showed that very few PPI indexes exhibited signs of non-response bias and the ones that did were affected by unusually high non-response in very specific size classes. These findings did not identify strong evidence of non-response bias in PPI indexes for the industries and years that were analyzed. Still the PPI program will continue to research methodologies capable of improving the response rate. It has implemented methods of monitoring response rate and contacting respondents who have ceased to participate. Non-response rates can be kept to a minimum with careful selection of industries to be resampled, improvements to sampling methodology, and concerted non-response follow-up efforts.

#### **4. Testing Procedures or Plans**

The PPI is not currently planning any procedural or methods tests requiring OMB approval.

#### **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.

#### **PPI Methodology References**

The methodology of the PPI has been documented in numerous papers and articles covering a broad spectrum of topics ranging from price theory and program concepts to actual data collection methodology. A list of references includes:

Bureau of Labor Statistics, *BLS Handbook of Methods*, U.S. Department of Labor.

Producer Price Index  
1220-0008

Available at <https://www.bls.gov/opub/hom/pdf/homch14.pdf> Chapter 14

PPI Methodology Reports:

<https://www.bls.gov/ppi/methodology.htm>

The Bureau of Labor Statistics in the Monthly Labor Review has published additional articles on specific PPI topics. All the articles can be accessed from this web page:

<http://www.bls.gov/ppi/ppimlr.htm> .

Additional articles related to PPI can be found at:

<http://www.bls.gov/opub/btn/archive/home.htm>