U.S. Energy Information Administration

Office of Energy Statistics

Office of Oil, Gas, and Coal Supply Statistics

Supporting Statement for Survey Clearance

U.S. Energy Information Administration Survey:

FORM EIA-914, Monthly Crude Oil, Lease Condensate, and Natural Gas Production Report

OMB No. 1905-0205

Part B:

Statistical Methods

Original Date: August 2014

Revised Date: November 2014

Table of Contents

B STATISTICAL METHODS....1

B.1.1 Purpose of Survey....1

B.1.2 Respondent Universe....1

B.2 Statistical Sampling, Imputation, and Estimation Procedures....2

B.2.1 Sampling....2

B.2.2 Estimating....3

B.2.3 Frame Maintenance....4

B.2.4 Efforts to Reduce Total Survey Error....4

B.3 Maximizing Response Rates....5

B.4 Testing Procedures....5

B.5 Statistical Consultations....5

B STATISTICAL METHODS

B.1.1 Purpose of Survey

The purpose of the proposed expansion of the Form EIA-914 survey, *Monthly Natural Gas Production Report*, is to collect more reliable and timely monthly natural gas, crude oil, and lease condensate production information for the lower 48 states. The goal is to publish accurate information not more than 60 days after the close of a report month.

Expanding coverage of natural gas production will allow more relevant data to be made available in a timely fashion, and new collection of crude oil and lease condensate production (reported together, and including API gravity categories) will enable tracking of changing domestic oil quality and will inform the discussion of crude oil export policies. Crude oil and lease condensate will be collected as a single volume. The proposed Form EIA-914 expansion will separate production reporting for Arkansas, California, Colorado, Kansas, Montana, North Dakota, Ohio, Pennsylvania, Utah, and West Virginia in addition to the current reporting for Texas (including State Offshore), Louisiana (including State Offshore), Oklahoma, New Mexico, Wyoming, Gulf of Mexico Federal Offshore, and Other States (defined as all remaining states, excluding Alaska).

For the current Form EIA-914, natural gas production estimates are replaced in EIA's publications and website with state reported production volumes when the state volumes are deemed complete. Similarly, for the expanded EIA-914, the crude oil and lease condensate production estimates and the gas production estimates will be replaced with state reported volumes when state reported volumes are deemed complete. EIA has procedures in place to determine when state reported volumes are complete, which for some states requires a period of up to two years from the reporting month. Oil and gas production data are included in numerous EIA publications, including the Petroleum Supply Annual (PSA), Petroleum Supply Monthly (PSM), Natural Gas Annual (NGA), Natural Gas Monthly (NGM), and the EIA website.

B.1.2 Respondent Universe

The universe of the expanded Form EIA-914 consists of all operators of wells in the lower 48 United States that are producing natural gas, crude oil, and/or lease condensate, including offshore wells. A well operator producing natural gas, crude oil, and/or lease condensate forms the responding unit.

A monthly cut-off sample of well operators is designed to give the targeted coverage in all areas with the minimum number of operators while reducing sample turnover from month to month. During 2013, EIA met this quality goal with about 240 respondents for the current survey.

EIA expects that expanding from seven to 17 states/areas and sampling oil and lease condensate production will increase the sample size from 243 to no more than 600 respondents. The cut-off sample is designed to provide about 90 percent coverage of all products at the lower-48 level for natural gas, crude oil and lease condensate and an adequate percent coverage in each geographic area for which production data are collected.

The cut-off sample operators are selected based on their natural gas and oil production obtained from a current commercially available data source, "Drilling Info" (DI), supplemented with data collected on the

Form EIA-23L, *Annual Survey of Domestic Oil and Gas Reserves*. These sources provide readily consolidated data that is recent and accurate.

Currently, the EIA-914 frame consists of about 13,000 active operators of crude oil and gas wells compiled from the DI database and Form EIA-23 responses.

B.2 Statistical Sampling, Imputation, and Estimation Procedures

B.2.1 Sampling

The DI database supplemented with Form EIA-23 data is used for both the sampling and estimation processes. DI acquires well or lease level oil and natural gas information from state agencies, places it in their own database format, performs some data standardization and validation, and sells the data. A new DI database is acquired by EIA at the end of every month.

DI data for four of the smaller producing states are missing or inadequate. For Illinois, Indiana, Kentucky, and Tennessee, annual production data from the Form EIA-23L survey are used to supplement the DI database. Hereafter, references to DI data include supplemental data from the Form EIA-23 survey for these four states. These data are used to determine a different cutoff sample for each state/area designed to achieve a roughly 85 percent sample coverage of that area's oil and natural gas production volumes. The cutoff production rates will be re-determined, annually, to maintain the target of 85 percent coverage.

DI data for the most recent reporting months are usually incomplete. The months used for sampling are determined by how much of a lag there is in the DI data before they become complete. A different lag is used for every state/area and for oil and gas. The process used to determine the lag time to complete data will be described in the published EIA-914 Methodology document.

For larger operators not in the sample, the production is examined for the most current four-month period where DI data are relatively complete for each operator. If a non-sampled operator is observed to produce above the cutoff for four consecutive months, that company is added to the sample. If a sampled company reports production on the Form EIA-914 below the cutoff in every area in which it produces, for six consecutive months, it is dropped from the sample. These procedures maintain high sample coverage of the population, and minimize frequent adding and dropping of operators.

Rationale for 85 percent sample coverage target

The 85 percent target was selected both to mimic the sampling rate of the previous EIA-914 survey, and to guard against model failure. In general, relative standard errors are within acceptable ranges at an 85 percent sample volume coverage. Of more practical concern is that in some specific cases, large operators can behave very differently than small operators. This occurred in Louisiana during 2009 when the Haynesville Shale was developed. Model failure in a cutoff sample occurs when the sampled companies do not represent the non-sampled companies according to the proposed model; the non-sampled companies behave differently than the sampled companies. By using a high sample coverage, the possible impact of model failure is limited.

The problem of model failure could normally be solved by combining a certainty stratum with a probability-proportional-to-size sampling or other random sampling scheme of the smaller companies. Because the sampling frame is in a constant state of flux due to operator acquisitions

and mergers, sample weights calculated on a monthly basis are prone to error. Using the classical ratio estimator, this sort of behavior causes much less fluctuation in estimates.

Another way of solving the model failure problem is through stratification of variables based on well characteristics, such as formation type (shale, conventional, etc.). While the frame has such information, it is often reported inconsistently and sometimes incorrectly. Also, to be effective the respondent would have to provide this information in their survey response since an individual operator may produce large volumes from multiple wells in multiple formations. Not only would this increase burden on the respondent, in EIA's experience with the EIA-23 (the Annual Oil and Gas Reserves Report), respondents do not provide this information reliably enough to clean the data within the necessary turnaround time.

Major producing states such as Kansas and Oklahoma have higher numbers of small oil producers than other major states. Smaller producing states have mostly smaller oil producers. States composed primarily of small operators and without potential for dramatic growth are generally much less likely to suffer from model failure. As a result, the below target sample coverage expected for liquids in states like Kansas and Oklahoma is acceptable.

Rationale for 500 barrel per day minimum threshold

In addition to the 85 percent sample coverage criteria used to determine the oil and gas production cutoff values, a minimum crude oil cutoff rate of 500 barrels per day (bpd) is used as an additional constraint to limit the number of smaller operators sampled. This is needed because some liquids producing states, such as Kansas, are composed of an inordinate number of small operators, and achieving an 85 percent sample coverage would create a much larger sample in a region that EIA does not expect model failure to be an issue. Application of the 500 bpd minimum oil cutoff rate will yield less than 85 percent coverage in some smaller states, but also reduces the total sample size by about 500 operators. EIA believes application of the oil cutoff rate creates an acceptable compromise between sample coverage and reduced burden on smaller operators.

The minimum cutoff rate of 500 b/d was selected in part because this was used effectively and without controversy on the Form EIA-23. EIA also considered what impact various minimum cutoffs would have on the sample and resulting data quality. Sample coverage worsens at increasing cutoff levels (see Table of Percent Coverage by State below). With no minimum cutoff, coverage across all states meets the EIA goal of 85 percent, but sample size increases considerably, thereby increasing burden. Boosting the minimum threshold to 500 bpd, 1,000 bpd, and finally 5,000 bpd succeeds in lowering the sample size, but generates progressively worse coverage and less reliable estimates.

Percent Coverage by State for Sampled Oil Production				
	Minimum Cutoff Applied (bpd)			
State	**0**	**500**	**1,000**	**5,000**
Arkansas	86%	41 %	30%	27%
California	93%	93 %	92%	91%

Colorado	90%	85%	85%	79%
Federal Gulf of Mexico	95%	94%	94%	93%
Kansas	86%	51%	42%	29%
Louisiana	87%	79%	73%	55%
Montana	92%	88%	88%	85%
New Mexico	92%	92%	92%	88%
North Dakota	95%	95%	95%	94%
Ohio	87%	69%	69%	69%
Oklahoma	86%	61%	61%	53%
Other States	90%	74%	74%	60%
Pennsylvania	87%	50%	50%	50%
Texas	91%	89%	89%	81%
Utah	95%	96%	96%	86%
West Virginia	84%	76%	76%	76%
Wyoming	89%	87%	87%	77%
Lower 48	**91%**	**90%**	**89%**	**84%**
Sample Size	**982**	**517**	**470**	**330**

The proposed expansion of the Form EIA-914 will utilize 34 cutoff samples, one for each of the 17 states/areas for crude oil and lease condensate production, and another 17 cutoff samples for each of the states/areas for natural gas production. If an operator is sampled for one state/area for either crude oil or natural gas, then it must report natural gas, crude oil and lease condensate for all states/areas where it has production, regardless of volume in those areas. This sample methodology is expected to increase the sampled set of operators from about 240 to fewer than 600 companies.

The proposed Form EIA-914 utilizes samples for the 16 states for which crude oil, lease condensate, and natural gas production respondents must provide information. The remaining 17 states (the “Other States” group) that have crude oil, lease condensate, or natural gas production will also be sampled, but collectively rather than individually.

B.2.2 Estimating

The target population of the proposed expansion of the Form EIA-914 is all active crude oil and lease condensate, and natural gas producers. There will be an estimation procedure for members of the target population not sampled for this data collection.

Estimation will be performed separately for natural gas and liquids production for each of the 17 states/areas using the Simple Ratio (SR) method, which multiplies the current survey reported production by a ratio. The SR ratio is an area's total production in an earlier 6-month period as found in the DI database, to the area's production from the current sample of operators in the same period, also found in the DI database. A six-month ratio is used to avoid potential volatility from using a ratio of a single month's production. The DI data are typically incomplete for more recent months. The months used in calculating the ratio are determined by how much of a lag there is in the DI data before they become sufficiently complete. A different lag is used for each state/area and for liquids and natural gas. Estimates for the survey month are computed using the following formula for each area:

$$\widehat{TP}_m = S_m \left(\frac{\sum_{i=L}^{L+5} TP_{m-i}}{\sum_{i=L}^{L+5} SP_{m-i}} \right)$$

where:

$\widehat{TP}_m = ¿$ Estimated total production for the state/area in month 'm'

$S_m = ¿$ Total survey reported production for the state/area in month 'm'

L = The DI lag for the particular state/area

$TP_i = ¿$ Total production for the area in month 'i' from the DI database

$SP_i = ¿$ Production of the current sample of operators in month 'i' from the DI database

Errors can occur in the estimation process when producing properties are transferred between operators, and this transfer is not reflected in both the DI database and the Form EIA-914. This can be caused by mergers and acquisitions. EIA monitors these events by following trade journals and news releases and by following up with respondents whose reported production is substantially different from what is expected given their production history. When transfers and other changes occur, EIA adjusts either the DI data or the survey data, as appropriate, to synchronize reporting until the two are in sync again.

All estimates will be based on the simple ratio estimate method with the exception of "Other States," which is currently determined for natural gas producers by applying the ratio of total gross natural gas production published in the NGA to the EIA-914 surveyed production for the most recent available year. The numbers from the NGA come from state-maintained webpages. This is used instead of the simple ratio method because the frame for the "Other States" group is less complete. If the frame coverage in the other states improves it could be estimated using the simple ratio estimate method as well. For the expanded EIA-914 the "Other States" is roughly 3 percent of both oil and gas production in the lower 48 states. An estimation method similar to the current "Other States" estimation process is anticipated, using the NGA data for gas and the PSA data for oil. Initially, the "Other States" volumes may be published as reported with no estimation until the adequacy of the "Other States" frame can be determined.

Imputation for Item Non-Response

Currently, Form EIA-914 experiences nearly 100 percent unit and item level response rates. However, EIA anticipates that item non-responses are more likely to occur in the expanded Form EIA-914, given additional reporting categories. In the event that item level non-responses cannot be resolved within a survey cycle, EIA intends to use the previous three months' data to impute current month data for a non-respondent. EIA will replace the imputed data with respondent data, once reported to EIA. If any of the three previous months' survey data are deemed not analogous to the current missing reporting month's data (i.e., being influenced by maintenance, weather event, power outage, etc.) that previous month's data would not be used in the imputation process.

B.2.3 Frame Maintenance

Since the inception of the Form EIA-914 in 2005, EIA has maintained a natural gas operator frame of the nation's natural gas producers. EIA uses various sources of information to maintain the completeness of the frame, including trade press, other data providers, and state regulatory agencies. EIA plans to maintain the natural gas, crude oil, and lease condensate frame by following similar practices.

EIA continuously reviews all available information to adjust the survey frame for the Form EIA-914 data collection for births, deaths, mergers, and company information changes.

B.2.4 Efforts to Reduce Total Survey Error

Frame Coverage Errors

In states with stricter reporting requirements, DI obtains more complete data in a timely fashion; frame coverage error is expected to be negligible for those states. These states make up the majority of U.S. production. However, in a minority of generally smaller producing states, data are less timely and less complete, requiring continuing effort on the part of EIA to research companies that may be operating in these areas to ensure the best possible frame coverage.

Reporting Errors and Data Processing Errors

The proposed Form EIA-914 questionnaire has been carefully developed and tested with potential respondents to minimize reporting errors. The survey instrument includes a detailed set of instructions for filing data, subject to a common set of definitions similar to those already used by the industry. EIA continually enhances data validation/editing software to detect probable reporting errors and flag them for resolution by analysts, either through confirmation of the data by the respondent or submission of amendments to the previously-filed data. Data processing errors by EIA are detected by the same software and are minimized through direct data collection from respondents.

Estimation Errors by Respondents

EIA will use agency-developed software to detect Form EIA-914 responses that appear out of the ordinary and are candidates for further investigation. Software will be used to flag data items that are inconsistent or out of their normal range, which will be further investigated by EIA analysts. EIA analysts will contact the respondents to resolve related questions.

Revisions

Revisions to aggregated data from respondents submitting revised data for prior months or late data for the current month are published according to the Form EIA-914 Revision Policy. This revision policy will be published along with the EIA-914 Methodology documents. Publication revision thresholds will be

determined for each state/area and for oil and gas. In addition, crude oil and lease condensate (combined), and natural gas production estimates will be replaced with state reported volumes when state reported volumes are deemed complete.

Respondents are required to resubmit data when their production estimates change by more than 1,000 barrels of crude oil and lease condensate in a state, or by more than 150 million cubic feet of natural gas in a state, and this requirement is detailed in the survey instructions.

B.3 Maximizing Response Rates

EIA will use standard procedures to administer the data collections for the proposed Form EIA-914. An introductory letter signed by the responsible EIA official is sent to each company that will be a respondent to the Form EIA-914. Follow-up procedures for non-response consist of an email message or a reminder phone call (for those not using email) to all companies that do not return a completed survey form by the due date. This initial re-contact is followed by repeated email messages, letters, and phone calls until an appropriate response is received.

B.4 Testing Procedures

The Form EIA-914 collection instrument questions were developed by starting with questions from the currently approved version of Form EIA-914, which the expanded Form EIA-914 is to replace. A team of EIA crude oil and lease condensate, and natural gas production analysts and survey experts collaborated to develop and test the new (expanded) Form EIA-914.

EIA conducted a number of cognitive interviews with producing companies to understand respondents' response processes and any potential barriers to survey form completion, including data availability and the timing of survey cycles. Cognitive interviews with ten companies were conducted in Texas during May 2014. As a result of the first round findings and subsequent revision of the form, a second round of cognitive interviews (five) was conducted in Pennsylvania and West Virginia during July 2014.

B.5 Statistical Consultations

For additional information concerning this data collection, please contact Jeff Little at (202) 586-6284, or jeff.little@eia.gov.

For information concerning this request for OMB approval, please contact the agency Forms Clearance Officer, Alethea Jennings, at (202) 586-5879, or alethea.jennings@eia.gov.