U.S. Energy Information Administration

Office of Energy Statistics

Office of Energy Consumption and Efficiency Statistics

Supporting Statement for Survey Clearance

Manufacturing Energy Consumption Survey

OMB No. 1905-0169

Statistical Methodology

Part B

Original Date: July 2014

Table of Contents

Collection of Information by Employing Statistical Methods.....1

1. Description of the Survey Plan.....1
2. Sampling Methodology and Eswtimation Procedures.....1
3. Maximizing the Response Rate.....4
4. Tests of Procedures.....4
5. Statistical Consultations.....5

B. Collection of Information by Employing Statistical Methods

1. Description of the Survey Plan

The 2014 MECS survey plan will largely follow the 2010 plan. The 2014 MECS will use the 2012 Economic Census - Manufacturing (ECM) mail file along with updates for births and deaths to identify establishments eligible for MECS. The MECS will present estimates for U.S. manufacturing industries at the National and Census Region levels.

The U.S. Census Bureau will again conduct the fieldwork for the survey, acting as EIA's data collection agent. The target population for the 2014 MECS will be comparable to that of the previous MECS, a restricted universe of manufacturing establishments above a minimum size threshold established for all NAICS industries in the manufacturing sector.

An Internet-based electronic reporting method will be available for the 2014 MECS. The Census Bureau will mail a letter along with the questionnaire. For those establishments that have Internet access the letter will give detailed instructions on how to complete the questionnaire through the Internet. However, for those establishments without Internet access or who choose not to use the Internet data collection method, the mailed questionnaire may be completed and sent back to the Census Bureau. Non-response follow-up procedures, as described in Section 3 below, will be similar to those used in 2010, which are performed to maximize response rate. Data processing and editing will be aided by modernized, automated data entry and storage equipment and editing procedures. Economic data from the 2014 Annual Survey of Manufacturers (ASM), where appropriate, will again be added to the MECS data file to provide extra analytic capability with no added burden on respondents. For non-ASM cases, economic data will have to be imputed or ratio adjusted from the 2012 ECM.

2. Sampling Methodology and Estimation Procedures

The 2014 MECS sample will be approximately 15,500 establishments. The frame for the MECS will be based largely on the NAICS classifications of the 2012 ECM mail file with updates through the three-year period. The 2012 ECM mail file was approximately 216,000 manufacturing establishments, and accounted for the top 97 to 98 percent of the total payroll in the manufacturing sector.

The 2014 MECS sample design will be stratified for NAICS industry with no provision to sample according to Census Region. However, industry-by-region data will still be published as in previous MECS. The major portion of the MECS sample will be devoted to the strata defined by the 3-digit NAICS subsectors that comprise manufacturing and 30 to 40 4-digit industry groups and 6-digit industries, all at the Census Region level. The industry groups and sub-industries will be chosen based upon their fuel consumption, economic output, feedstock usage, other programmatic interest, and

carbon dioxide intensity. In 2010, the MECS sample had 29 industries that entered with certainty (i.e., all establishments listed for those industries were selected into the MECS sample). Target coefficients of variation (CV) for the major energy sources for industry strata are set between zero (for certainty industries) to nine percent, depending on the NAICS level (six, four, or three digit) as well as other factors such as energy intensity.

The measure of size (MOS) is used to assign overall probabilities of MECS selection to each establishment, excluding the MECS certainty strata. The 2014 MECS MOS will be the sum of annual cost of fuel and cost of electricity (CF). These data come from the ASM, the ECM, or are imputed. The ASM is a sample survey of approximately 50,000 manufacturing establishments and measures many of the same economic variables that the ECM does, including cost of fuels. The 2012 ECM CF will be used to determine the 2014 MOS. For cases that are birth establishments since the 2012 ECM, a value will be imputed using 2012 payroll or employment data and ratio-adjusting the payroll by the median 3-digit NAICS 2012 ECM CF to 2012 payroll. Imputation is used when necessary for all establishments that have missing cost of fuels and electricity data. The method is usually a ratio estimator using establishment payroll data. The 2012 ECM mail file with adjustments for births and deaths will be used to determine what establishments would be eligible for MECS sampling.

The MECS uses a probability proportionate to size (PPS) approach to assigning inclusion probabilities to manufacturing establishments in the strata. However, the MOS is not highly correlated with measured values of energy consumed as a feedstock and certain energy sources used as a fuel. To address this, the 2014 MECS will select with certainty establishments in the frame known to be significant consumers in these areas.

After the inclusion probabilities are assigned, a modified version of a sample selection algorithm presented by Pareto[1] is employed to obtain the 2014 MECS sample. This procedure, first used for selecting the MECS sample in 1998, ensures that the sample in each stratum will be of a fixed, predetermined size. This benefits the MECS by preventing cost overruns due to the actual sample size being larger than was expected. Additionally, this selection method was used instead of Tillé sampling, the selection method used in 2010, because the variance is simple and easy to calculate and does not involve joint probabilities, and Pareto is easy to implement, as well as the ease of transferring knowledge to maintain the sampling program if in the future certain modifications are desired.

The statistic reported for each stratum is an estimator of the total of some measure of energy consumption (e.g. the total natural gas, in billion cubic feet, for the Northeast primary metals industry). Modified Horvitz-Thompson estimators are used for estimating totals. For each stratum, the estimates are of the following form:

$$\hat{Y}_s = \sum_{i=1}^{n_s} w_i^{a\hat{\phi}} \cdot (y_i)$$

[1] It was introduced in two articles by Bengt Rosén [Journal of Statistical Planning and Inference 62 (1997) pp. 135-158 and pp. 159-191]. It is also referenced by Särndal and Lundström (Estimation in Surveys with Nonresponse, Wiley 2005 p. 31).

where:

yi , *i* = 1 : *ns* are the values of the consumption measure for the *ns* establishments in stratum *s* that were selected for the sample and responded.

w_i^{adj} , *i* = 1 : *ns* are the final sampling weights for the *ns* establishments in stratum *s* that were selected for the sample and responded.

$\hat{Y}_s$ is an estimate for $Y_s = \sum_{i=1}^{N_s} y_i$ which is the actual total for stratum *s*

Population totals will be estimated from 2014 MECS data by summing the nonresponse-adjusted weighted sample data. For the 2010 MECS, the nonresponse adjustments were done separately for certainty and noncertainty establishments within the cells. That practice will likely be continued for 2014.

The MECS provides estimates for the number of establishments consuming a particular type of fuel (e.g. coal) by NAICS code at the national level or for the number of establishments participating in an energy management program. An adjusted Horvitz-Thompson estimator is used for this. For industry group s and fuel group f, the estimate for the number of establishments is:

$$\hat{U}_{s,f} = \frac{N_s^{cert}}{\hat{N}_s^{cert}} \sum_{i \in s_r^c} u_{i,f} \cdot w_i^{adj} + \frac{N_s^{noncert}}{\hat{N}_s^{noncert}} \sum_{i \in s_r^{nc}} u_{i,f} \cdot w_i^{adj}$$

$$u_{i,f} = \begin{cases} 1 & \text{if establishment i in the sample for industry s consumed fuel type f} \\ 0 & \text{otherwise} \end{cases}$$

where:

w_i^{adj} is the poststratified sample weight for the i^{th} establishment in the sample.

s_r^c / s_r^{nc} is the set of certainty/noncertainty establishments that were selected and responded to the MECS in industry group *s*.

Ns is the number of establishments in the MECS frame which fall into the certainty/noncertainty portion of industry group *s*. This control total comes from the ECM.

$\hat{N}_s = \sum_{i=1}^{n_s} w_i^{adj}$

= sum of the poststratified nonresponse-adjusted weights for all responding establishments in the certainty/noncertainty portion of industry group *S*.

For a more detailed and in-depth explanation of the MECS methodology please read about the methodology on our website: http://www.eia.gov/consumption/manufacturing/data/2002/index.cfm?view=methodology. Although this explains the 2002 methodology, it will be the same methodology employed in the 2014 MECS.

3. Maximizing the Response Rate

The weighted coverage rate (ratio of the total weighted MOS of the responders to the total weighted MOS in the cell) for the 2010 MECS was 88 percent. By using the Census Bureau as the data-collection agent, the survey is not only under the data-collection authority of the DOE mandate, but also has the confidentiality protection that Title 13, Section 9, of the U.S. Code confers on surveys conducted by the Census Bureau.

The Census Bureau has developed standard procedures for non-response follow-up and imputation. The Bureau will conduct two mail follow-ups, as well as a telephone follow-up of the largest delinquents. Smaller delinquents will receive only the two mail follow-ups. An enhancement to the telephone nonresponse follow-up instituted for the 2002 MECS will be continued for the 2014 MECS. In that procedure, follow-up was done to maximize response of establishments that have the greatest MOS in the adjustment cells. The procedure established target coverage rates (ratio of the total MOS of the responders to the total MOS in the cell) for the follow-up based on the overall importance of the cell.

4. Tests of Procedures

This will be the ninth time the MECS will be conducted. After consideration of past results and outside consultations (as described in Section A-3), the questionnaire has undergone substantial redesign in the method of collection. The effectiveness of these changes was analyzed during the editing and review of the 2010 MECS data. The respondents actively supported the changes during the post-survey interviews of respondents in 2012. The few changes to the instrument for 2014 target opportunities for increased data reporting frequency, improved data quality, and reporting efficiency.

The sampling procedure was performed successfully for the 2010 MECS. Tests will be run, where possible, to determine the best way to adjust the MOS to 2012 values when ASM data are not available.

In carrying out the 2014 MECS, the Census Bureau will use fieldwork procedures similar to those routinely used for the ASM and that were used for the 2010 MECS. No further tests of procedures are planned.

5. Statistical Consultations

The U.S. Bureau of the Census is selecting the MECS sample under the supervision of David Kinyon. Mr. Kinyon can be reached at (301) 763-7209. The overall administrative responsibility for MECS at the Census Bureau rests with Nick Orsini. Mr. Orsini can be reached at (301) 763-6959. The principal EIA official consulted about the sample design is Tom Lorenz, Survey Statistician, of the Office of Energy Consumption and Efficiency Statistics within the Office of Energy Statistics. Mr. Lorenz can be reached at (202) 586-3442. The EIA Agency Clearance Officer is Alethea Jennings, 202-586-5879.