SUPPORTING STATEMENT PART B
FOR FORMS EIA-871A-J
COMMERCIAL BUILDINGS ENERGY CONSUMPTION SURVEY

OMB NO: 1905-0145

B-2. The Sampling Methodology and Estimation Procedures

Sampling Methodology

The 2012 CBECS design provides estimates of baseline energy consumption and related building characteristics for the universe of commercial buildings in the United States. The survey results will be presented at the national level, the four Census regions, and in instances where the data allow, at the Census division level, as well as by other characteristics such as climate zone, building type, total square footage, number of floors, and year constructed. The top priority levels of aggregation are the National and Census region estimates.

Similar aggregations of data were produced from the 1979, 1983, 1986, 1989, 1992, 1995, 1999 and 2003 CBECS. The 2007 CBECS sample design did not yield statistically valid results. For the 2003 CBECS, the sample of 5,215 completed building interviews resulted in a relative standard error (RSE) of approximately 4 percent for national estimates of total energy consumption. At the regional level, energy consumption RSEs were between 5 and 11 percent, depending on the region. For energy expenditures at the national and regional levels, the RSEs were similar to the RSEs for total consumption (See Tables C3A-C6A: Consumption and Expenditures RSE Tables: http://www.eia.gov/emeu/cbecs/cbecs2003/detailed_tables_2003/2003rsetables.html)

Similar to previous CBECS, there are two parts to the 2012 CBECS sample design, a multistage area probability sample and a supplemental sample of special buildings.

Multistage Area Probability Frame

The 2012 CBECS is based on a sample frame and design that was new for the 2003 CBECS. The 2003 design is described below. Previous CBECS were based on a 1986 sample design that was updated for new construction for each subsequent survey cycle. The general approach of the multistage area probability sample design is to sample and subsample successively smaller geographic areas until it is feasible to list all commercial buildings in the selected areas.

Primary Sampling Unit (PSU): In the multistage area sample portion of the 2003 CBECS design, the approximate 3,100 counties and independent cities in the 50 States and the District of Columbia are grouped into 687 PSUs. The boundaries of these PSUs are contiguous with the boundaries of counties, and independent cities. Of the 687 PSUs in the frame, 108 are selected for the sample PSUs.

The PSUs in the frame are grouped into 65 strata that are relatively homogeneous based on geographic characteristics such as similar climate and topography. These strata are designed not for estimation purposes, but as representative homogeneous groupings from which PSUs are selected. Because of this rationale, a given stratum may include geographic areas that are not contiguous.

Twenty of the 65 strata contained only one PSU and were sampled with certainty. The remaining 45 noncertainty strata contained multiple PSUs. In 43 of these strata, two PSUs were selected using probabilities proportional to a measure of size that relates to commercial activity (in the last 2 strata, only one PSU was selected). These selections resulted in 88 noncertainty PSUs and 20 certainty PSUs.

Secondary Sampling Unit (SSU): The next step in the multistage area sampling procedure consists of selecting subareas within each sampled PSU. Within the 108 sample PSUs, 7,031 SSUs were constructed by combining 23,117 Census tracts. Of these, 511 SSUs were selected in the main sample and 220 were in a reserve sample to be used only if one of the main SSUs had no buildings or if there were a shortage of listed buildings at the PSU level. The boundaries of the SSUs coincide with those of the Census tracts available from the Bureau of the Census.

Buildings: The final step consists of selecting buildings within each sampled segment. A segment is the same as an SSU in most cases. In geographically large SSUs, segments are smaller areas of the SSU that are broken up, or "chunked," and subsampled so that these areas are feasible to list. A field listing was conducted in 515 segments (511 main sample and 4 reserve sample segments) that produced a listing of all commercial buildings, their size and type. Buildings are selected for inclusion in the sample from these listings. Sampling rates within each size and building use category are calculated with an optimization formula that minimizes the variance of total estimated energy consumption.

List Frame of Special Buildings

It is desirable to sample large buildings at a higher rate than small buildings because of the relatively large amount of energy that is consumed in large buildings. However, the area sampling procedure cannot provide an optimally efficient mix of large and small buildings unless an exceedingly large number of buildings are listed in order to guarantee that a sufficient number of very large buildings would be available for sampling.

To compensate for this inefficiency of the area sample, special lists of government buildings, schools, hospitals, airports, and other large buildings are prepared for the sample PSUs. To complete the full 2012 sample, buildings will be sampled from these lists at the PSU level and then combined with those selected by the area sampling procedure after checking the frames for duplication.

The list frame consists of: U.S. Government buildings that are at least 200,000 square feet (from

the General Services Administration Real Property Inventory); four-year colleges and universities that are estimated to have at least 1,000,000 square feet of floorspace (from the Integrated Postsecondary Education Data System); hospitals that are estimated to have at least 200,000 square feet of floorspace (purchased from the IMS Healthcare Market Index); a list of airports (from the Federal Aviation Administration); and a list of large buildings over 200,000 square feet (from the Common Premise Locations (CPL) file purchased from Dun & Bradstreet). The list of airports is new to the 2012 CBECS; previous versions of the other lists were used in the 2003 CBECS.

Update for New Construction

The new construction is covered by several different sources in the 2012 CBECS sample design. In the 20 SR PSUs, 61 new segments were selected. In addition, a total of 196 segments were selected from 43 new PSUs that were selected from the NSR strata. All of these 257 new segments will be listed and both old and new construction buildings will be sampled with rates to achieve the desired building overall selection probabilities. In addition to providing new construction coverage, these new PSUs and segments will help achieve the increased sample size. All supplemental list frames for large buildings were created using the latest available lists to provide highest coverage of new large buildings. The new buildings from these lists will be sampled to achieve the desired building sampling rates.

In addition, a 10% sample of update segments was drawn from the 2003 sample segments, where only these update segments will be dependent-listed to identify changes to the 2003 list. The dependent-listing procedures will use the 2003 listings as a starting point for field listers to identify changes since 2003, specifically new and demolished buildings. A new sample of buildings will be drawn from the listings in the update and non-update segments. In the update segments, the buildings will be stratified by "old construction" (buildings confirmed to exist from dependent relisting) and "new construction" (buildings added to the list in 2012). In the old construction strata, the 2003 sampling rates will be applied (with further optimization based on lessons learned from the 2003 CBECS, such as non-response rates). In the new construction strata, the new listings will be sampled with higher rates to bring their overall probability of selection in line with old construction listings. The update segments were selected with probability proportional to size (PPS) with the segment weight as the measure of size. This selection procedure will help to achieve the desired overall building selection probabilities for the new construction avoiding small segment selection probabilities.

Sample Size

Approximately 12,340 buildings (including strip malls) and 1,700 mall establishments will be selected for the 2012 CBECS sample.

Assuming an 83 percent in-scope rate and an 82 percent response rate, the 12,340 buildings selected will yield the target of 8,400 completed building interviews. This target represents

about a 1.5 increase from the 2007 CBECS target of 5,625 complete building interviews. The sample is being increased mainly in order to improve the statistical precision of the estimates and reduce the number of suppressed cells in the published tables, revealing more estimates by building activity at lower levels of geography.

To arrive at the decision to increase the sample size by 1.5, EIA studied the effects of both a 1.5 times increase and a 2 times increase. To determine the effect of a doubled sample size on the published CBECS data, a simulated data set was created that had two copies of each sampled building from the 2003 data set. To reflect the higher probability of selection with a doubled sample, and to make the simulated estimates consistent with the 2003 published estimates, the weight of each copy was halved.

For variance estimation purposes, the usual jackknife procedure was used. Each simulated building remained in the same stratum but was randomly assigned to a group in the formation of psuedoreplicates. The random assignments were independent of a building's actual group in the variance estimation for the real 2003 CBECS.

For the "1.5 sample" data set, a similar procedure was followed, except each full sample weight was multiplied by 2/3, and a randomly selected half of the buildings were duplicated on the final data set.

The tables produced from these simulated data sets displayed results that were consistent with statistical theory, in that variances were lower and more table cells met the threshold for publication.

As an example, in a crosstabulation defined by nine Census divisions as column headers and 16 building activities as row categories, with the original sample size, there are 51 cells suppressed out of the total 144 cells (35% suppression). This same crosstabulation with the 1.5 sample yielded only 28 suppressed cells (19% suppression), while the double sample still yielded 24 suppressed cells (17% suppression).

As evidenced by this example, the 1.5 sample produced greater incremental statistical benefit (reduced suppression) than the second increment, the double sample. This observation, coupled with cost considerations, led EIA to a target sample size of about 8,400 buildings.

Estimation Procedures

Population Parameters: Estimates of population parameters will be prepared by weighting the values from the sample. For each sampled building, the base weight will be equal to the inverse of the probability of selecting the building. These base weights will be adjusted to account for nonresponse.

Variance Estimates: Variance estimates and relative standard error (RSE) estimates will be calculated by means of a replication technique such as the Jackknife or Balanced Repeated

Replications (BRR). The generalized RSEs for the statistics in the cells of most CBECS report tables will be expressed in separate RSE tables.

Data Collection

The 2012 CBECS will collect energy-related building characteristics and energy consumption and expenditures information from a sample of U.S. commercial buildings. After the sample of buildings are selected, introduction letters will be sent to give the building owners and managers advance notice that a representative of EIA will be contacting them. Then an interviewer will call or visit the building and provide the appropriate respondent with some background information about CBECS.

Included in this background CBECS information are two worksheets (Forms EIA-871G and H) for the respondent to complete prior to the interview. These worksheets will be formatted slightly differently depending on the interview type – for example, not all questions will appear on the versions to be used for the strip mall building and the mall establishment questionnaire forms, and the words "establishment" and "mall building" will be substituted for "building" on the mall building and the establishment questionnaire worksheets. Respondents to the mall building questionnaire will not require Form EIA-871H. The worksheets will help the interview to proceed more quickly and smoothly by allowing the respondent to obtain answers to several of the questions that may require some research prior to the interview.

At the time that they provide the background information to the buildings' owners, managers or tenants, the interviewer will also schedule an appointment for the CBECS interview, to be conducted either by telephone or in person. The majority of respondents will be administered Form EIA-871A, which is the Building Questionnaire. Data for strip shopping centers will be collected via several interviews with several different respondents within the selected centers. The strip mall manager or leasing agent will be administered Form EIA-871I, which will collect general information on the mall structure; several establishments within the mall will be subsampled and respondents from each of those businesses will be asked questions (Form EIA-871J) pertaining specifically to that establishment, such as operating characteristics, equipment used, and energy consumption.

In cases where the building respondent is unable to provide the consumption and expenditures data, the respondent will be asked to sign an authorization form (Form EIA-871B), giving their energy suppliers permission to release their billing records for the building. In these cases, the consumption and expenditures data will be collected from the energy suppliers on Forms EIA-871C-F. The energy supplier will be asked to provide the annual amount of energy sold, the annual cost of the energy, the billing dates, and other pertinent data specific to each type of energy source. The energy supplier data will be linked to the results of the Building Questionnaire for analysis purposes. Response to the energy suppliers forms (EIA-871C-F) is mandatory for the supplier.

The respondent burden on energy suppliers will be reduced because, for a substantial number of buildings, the energy consumption and expenditures data will be collected at the building level rather than at the supplier level. This will result in fewer forms the energy suppliers will have to process and increased overall efficiency.

In cases where energy suppliers must be contacted to obtain the consumption and expenditures data, EIA does not require these suppliers to transcribe the data onto the survey forms. EIA will accept responses from the energy suppliers in any format (such as a computer printout or spreadsheet) as long as all the necessary information is provided.

Similar to previous CBECS, the 2012 CBECS will not survey buildings or suppliers for the actual consumption of district chilled water, bottled gas (LPG or propane), wood or coal. Categorical estimates of bottled gas and wood consumption are collected from the building respondent, but these energy sources are usually purchased as-needed in small quantities, so neither would yield much information from an energy supplier.

B-3.   Maximizing the Response Rate

The data collection contractors will use procedures similar to those used on previous surveys to maximize the response rate for the 2012 CBECS with some additional effort. Over the years, the response rates to the CBECS have been quite high, but it is recognized that response rates have been declining over the years for surveys in general. The CBECS is no exception; response rates on the buildings survey have fallen from 91 percent in 1992 to 82 percent in 2003 and 81 percent in 2007. Great effort will be made in 2012 in an attempt to reverse that trend. Procedures that will be used to maximize the response rate make use of extensive pre-data collection procedures that include:

- Training interviewers in refusal aversion;
- Emphasis on the importance of the content of the survey, both to the interviewers and to the survey respondents;
- Training interviewers to locate a knowledgeable respondent;
- Mailing letters signed by a Department of Energy official to the respondent to inform them about the survey;
- Providing the identified respondent a packet of materials consisting of information about the survey, worksheets to assist them with the interview, and an introductory letter signed by the EIA Office Director;
- Including a well-designed brochure and an endorsement insert in the packet of materials that highlights the importance of the CBECS and why their participation is so vital;
- Setting an appointment to conduct the interview and providing the option to do so either in person or by telephone;
- Providing a Help Center with a toll-free number for the respondents to call with any questions or concerns, as well as the phone number of a contact at EIA;

- Following up with nonrespondents by various methods such as reminder letters (tailored for the reason for reluctance) and telephone calls for building owners or managers;
- Staffing two interviewers in most PSUs. This is advantageous to achieving high response rates because it allows reassignment of nonresponse cases among interviewers living in the same area, as well as depth to handle illness or attrition.

To adjust for nonresponding buildings, EIA adjusts the weights on responding buildings so that responding buildings represent not only nonsampled buildings but also nonrespondents.

B-4. Tests of the Procedures

EIA plans to conduct some targeted cognitive pretesting of new questions or procedures in late 2012. The survey contractor will also be performing systems integration testing in late 2012. This will be an operational test of the end-to-end systems involved in the survey.

B-5. Statistical Consultations

EIA has chosen SAIC as the prime survey contractor with Westat as the subcontractor for the CBECS building survey. At Westat, the program director is Robert Patchen. Mr. Patchen may be reached at (301) 610-5113 or by e-mail at bobpatchen@westat.com. The lead senior statistician from Westat is Huseyin Goksel. Mr. Goksel may be reached by e-mail at gokselh1@westat.com. The CBECS Survey Manager at EIA is Joelle Michaels. Ms. Michaels may be reached at (202) 586-8952 or by e-mail at joelle.michaels@eia.gov.

The agency point-of-contact regarding this clearance is Alethea Jennings. Ms. Jennings may be reached at (202) 586-5879 or by e-mail at alethea.jennings@eia.gov.