

# Estimating Characteristics of Strip Shopping Center Buildings in the 2012 CBECS

Jay Olsen<sup>1</sup>

<sup>1</sup>United States Energy Information Administration, 1000 Independence Avenue,  
Washington, DC 20585

## Abstract

The Commercial Buildings Energy Consumption Survey (CBECS) is a quadrennial survey of commercial buildings, such as offices, retail stores, and other nonresidential structures. The survey collects information on energy use in buildings and their characteristics that are related to energy use. Finding a respondent who is knowledgeable about the entire building has always been a particular challenge for strip shopping centers, as most of the people on site are store managers who can only give information about their own establishment. To help rectify this, the 2012 CBECS used a new procedure in which the interviewers compiled a roster of all of the establishments in a strip shopping center building, and conducted interviews at a randomly chosen subset of the stores. The new procedure will produce a description of the establishments that were listed on the roster but not interviewed by imputing based on similar establishments that were interviewed. The interviewed and imputed establishment characteristics will then be aggregated into a description of the entire strip shopping center building.

**Key Words:** Commercial Buildings Energy Consumption Survey, strip shopping centers, subsampling

## 1. Background

Every 4 years the United States Energy Information Administration (EIA) conducts a survey of commercial buildings in the U.S. The Commercial Buildings Energy Consumption Survey (CBECS) is a national sample survey that collects information on the stock of U.S. commercial buildings, including their energy-related building characteristics and energy usage data (consumption and expenditures). Commercial buildings include all buildings in which at least half of the floorspace is used for a purpose that is not residential, industrial, or agricultural.

The CBECS collects building characteristics information through a computer assisted personal interview with a respondent knowledgeable about the construction and operation of the building, including the types of heating, cooling, and other energy using equipment present in the building. Some commercial buildings have multiple tenants, each of which is an independently owned and operated business. In that case, it can be difficult to locate a respondent who has the requisite knowledge of the building as a whole. Experience with earlier CBECS surveys has shown that this problem is particularly acute with strip shopping centers, which are frequently owned by an off-site company and which typically have many smaller establishments.

To address this issue, the CBECS has been modified over the years to incorporate special procedures to collect information on strip shopping center buildings. As with other commercial buildings, an interview is done with a respondent who has some responsibility for the strip center building as a whole. However, this interview includes only a subset of the CBECS questionnaire, primarily focusing on information that pertains to the entire building, such as total size and building construction materials. Additional information about individual stores in the strip center is collected directly from respondents knowledgeable about those stores, and that information is used to estimate the characteristics of the strip center building as a whole. The development and testing of the latest refinement to these procedures incorporated into the 2012 version of the CBECS is detailed in the sections below.

## **2. Subsampling of Establishments at a Strip Shopping Center**

### **2.1 Establishment Roster**

In order to collect information for individual establishments, the exact stores, shops and restaurants at the site had to first be identified by the CBECS interviewer. In the 2012 CBECS, the interviewer was instructed to list all of the establishments at the strip center, and also to estimate the two characteristics that are most closely associated with energy use: the activity that occurs in the establishment, and the actual size of the establishment as measured in square feet of floorspace.

#### *2.1.1 Establishment Activity*

The interviewers were instructed to examine each establishment and categorize the use or activity by choosing the most appropriate description from a list of 12 categories. These categories were:

- Retail Store
- Food Sales (Grocery)
- Food Service (Restaurant)
- Dry Cleaner/Laundromat
- Movie Theater/Cinema
- Beauty Salon/Tanning Salon
- Fitness Center/Dance Studio
- Education/Child Care
- Office/Financial
- Medical Office (Doctor, Dentist)
- Vacant
- Other

Experience with previous CBECS has shown that these categories are sufficient to cover most strip center tenants, and also have sufficiently different energy consumption profiles to make it worthwhile to distinguish between them.

#### *2.1.2 Establishment Size*

The interviewers were also instructed to walk across the entire frontage of the strip shopping center and record the number of steps that they required to pass each store. This number was used as a proxy for the relative floorspace of the establishment. Of course, because each interviewer has their own length of stride, this number cannot be directly

converted into an actual linear measure in feet of the frontage of a store. In addition, since each store will have a depth that was not measured, a linear measure of the store frontage cannot be directly converted into an estimate of the floorspace area of the store.

However, this measure was primarily used to estimate the *relative* size of each establishment *within a single strip shopping center*. It was assumed that the interviewers' length of stride would be consistent throughout the process of measuring the frontage. Also, while it is not necessarily true that each store in a strip center has the same depth, it occurs frequently enough that the estimated store frontage could be used as an estimate of relative size for sampling purposes, as detailed in the section below.

## 2.2 Subsampling of Establishments

Because there were not sufficient resources to conduct a CBECS interview at every establishment at each strip shopping center, they were subsampled with probability proportionate to size (PPS), and interviews were conducted only at the stores chosen by this subsampling procedure.

Because the CBECS is primarily concerned with energy consumption in commercial buildings, the appropriate "size" for the PPS sampling was an estimate of the expected energy consumption of each establishment. To calculate this estimate, analysts used a data set of establishments collected in a previous CBECS to calculate the average energy intensity for each activity; that is, consumption of energy in British thermal units (Btu) per square foot for each activity.

$$Intensity_i = \frac{\sum_i Energy\ Consumption}{\sum_i Square\ footage}$$

For each activity  $i$ .

Then, for each establishment on the strip shopping center roster, relative predicted energy consumption was estimated as

$$Intensity \times frontage\ in\ paces$$

Again, this is not a direct estimate of energy consumption. There is a mismatch between the units in the denominator (square feet) and the measure of the store frontage (linear paces). However, this is sufficient for providing relative sizes within a strip center. The actual sampling algorithm involved was a SAS program that used PROC SURVEYSELECT.

## 3. Estimation of Strip Shopping Center Characteristics Using Subsampled Establishments

### 3.1 Establishment Square Footage

Once interviews were completed for each subsampled establishment, the collected information could be used to estimate characteristics for the establishments that were not sampled. The first and most important characteristic is the square footage of each establishment. This is estimated by calculating the square footage in the strip center that is unaccounted for by responding establishment, and distributing it proportionally among

the establishments that were not sampled or did not respond, according to their measured frontage from the establishment roster.

$$\text{Estimated size} = \text{frontage} \times \frac{\text{Total square footage} - \sum_j \text{square footage}}{\sum_k \text{frontage}}$$

Where

$j$  is the set of establishments within a strip center that had a reported square footage; and  $k$  is the set of establishments within a strip center that DID NOT have a reported square footage

### 3.2 Establishment Characteristics

Using the estimated establishment square footage calculated above, other characteristics of nonresponding establishments can be calculated. There are many different questions in the CBECS interview, and each one could potentially have a different method of estimation. Full details of the methodology used for each CBECS data item is beyond the scope of this paper, but the estimation method used for numeric variables can be explained in some detail.

There are many variables in the CBECS that reflect the total number of some item present in a commercial building. This can be the total number of workers, the total number of some type of energy using equipment (computers, refrigerators, etc.), or the total floorspace served by a particular type of energy using equipment (fluorescent lights, furnaces, etc.).

Characteristics of this type can be estimated by calculating an average presence per square foot for responding establishments, and multiplying this average by the estimated square footage for each nonresponding establishment to provide an estimate for those establishments. Before the average is calculated, the establishments are divided into subgroups that are related to the characteristic being estimated.

Some characteristics tend to be associated with the activity in the establishment (e.g., number of workers or computers), while some are associated with geographical information (e.g., type of heating or cooling equipment). “Geographical” can include establishments located within the same strip center, the same Census division, or the same Census region.

## 4. Testing of Estimation Method

### 4.1 Creating a Simulated Strip Shopping Center Data Set

In order to test the methodology described above, it was necessary to subsample from a data set where the characteristic to be estimated was already known for each strip center and each establishment. Since there were no data sets from previous CBECS where all of the required information was already known, the testing was conducted on a simulated data set that mimicked the characteristics of real-world strip centers as closely as possible. The simulated data set was created using the following procedure:

- A SAS data set was created with 400 observations, each one representing a “Strip shopping center.” At this stage, these “strip centers” were empty; that is, they had no information associated with them.
- Each “strip center” was populated with a variable number of “establishment” records. The number of establishments per strip center was random, but chosen from the same distribution that had been observed in the previous CBECS.
- Each establishment was given a random size and activity. Again, these were chosen from the same distribution that had been observed in the previous CBECS.
- Each establishment was assigned an intensity depending on the establishment’s activity. These were based on the average energy consumption per square foot for those activities observed in the previous CBECS.
- Each establishment was given an “energy consumption” variable, which was equal to the size times the assigned intensity, with some random variation. The variation was intended to prevent the resulting data set from being too uniform or deterministic, which would not provide a good test of the sampling methodology.

#### 4.2 Using the Simulated Data Set for Testing

The establishment subsampling procedure was performed on the simulated data set created above. The average consumption per square foot within subgroups defined by activity was calculated, and these averages were applied to the unsampled establishments in the data set. The estimated consumption for each establishment was this average times the establishment’s size.

The establishment level estimates were then used to calculate an estimate of the total consumption in the strip center. This estimate was the sum of actual consumption for sampled establishments in the strip center, plus the sum of estimated consumption for the unsampled establishments. This estimate was compared to the actual consumption for the strip center, as calculated by the sum of the actual consumption of all of the establishments in the mall. To test the results, a ratio was formed between the estimated and actual consumption of each strip center, and the distribution of the ratio was examined using PROC UNIVARIATE in SAS. The results are presented in Table 1 below.

**Table 1:** Quantiles of ratio between true strip center consumption and estimated strip center consumption

<i>Quantile</i>	<i>Estimate</i>
100% Max	3.347887
99%	2.882042
95%	1.851889
90%	1.538655
75% Q3	1.199932
50% Median	0.896549
25% Q1	0.50007
10%	0.205538
5%	0.12869
1%	0.042172
0% Min	0.0165363

## 5. Conclusions

The decision to subsample the establishments in strip shopping centers, and the division of the items CBECS questionnaire into separate interviews that are appropriate for mall managers and store managers had a number of effects. It has reduced the burden on those respondents, allowing the more rapid completion of those interviews. It has facilitated the task of interviewers and made it more straightforward to find appropriate respondents. It may also have increased the quality of the data collected, as respondents are asked questions about characteristics that they are more familiar with and have less need to estimate items for which they lack direct knowledge.

The data collected from the individual establishments can be used to estimate the characteristics of the strip shopping centers from which they were subsampled. These estimates are sufficiently similar to the true strip center values to be used in calculating nationwide estimates of the characteristics and energy consumption of strip center buildings.

## Acknowledgements

The author gratefully acknowledges Thomas Leckey, Director, Office of Energy Consumption & Efficiency Statistics (ECES) within the Energy Information Administration for supporting his attendance at the Joint Statistical Meetings and the presentation and dissemination of this paper.

The author also acknowledges Hiroaki Minato, leader of the End Use Consumption Team within ECES, for his guidance and assistance in the development of the methodology presented in this paper.

## References

Cochrane, William G., 1977. *Sampling Techniques, third edition*. New York, NY: John Wiley & Sons.