Independent Statistics & Analysis
U.S. Energy Information Administration

August 2016

Green Building Certification and the Commercial Buildings Energy Consumption Survey (CBECS)

Joelle Michaels, U.S. Energy Information Administration
Amanda Webb, The Pennsylvania State University

ABSTRACT

Questions about green building certification were introduced into the 2012 version of the Commercial Buildings Energy Consumption Survey (CBECS), resulting in the first nationally representative data on energy consumption and energy-related characteristics in both certified and non-certified commercial buildings in the United States. This study examines the validity of the survey responses to these questions and explores possible factors associated with false reporting. Survey responses were compared against records in the ENERGY STAR and LEED certified building directories, and statistical measures of validity were calculated. Correlation coefficients (ranging from 0.45 – 0.67) show only moderate strength of association between the survey responses and directory records, and sensitivities (ranging from 55% - 78%) and positive predictive values (ranging from 38% - 60%) were both lower than expected. These results indicate that CBECS respondents do not always correctly identify their building as certified, and the survey responses should not be considered valid. Evidence from logistic regression models and recordings of the interaction between the respondent and the interviewer suggests that several factors, including social desirability bias, are associated with false response. These findings have implications for the collection of this data in future CBECS cycles, as well as for the use of the current validated dataset.

Background

Since the previous iteration of the Commercial Buildings Energy Consumption Survey (CBECS) in 2003, there has been a dramatic increase in the number of certified green buildings in the United States. In 2003, fewer than 2,000 buildings had been certified under the U.S. EPA's popular ENERGY STAR program; by 2012 more than 20,000 had been certified[1]. The U.S. Green Building Council's widespread LEED certification program shows similar growth, with less than 100 buildings in the U.S. certified by 2003 and more than 11,000 certified by 2012[2]. This growth has been hastened by local mandatory

[1] Authors' cumulative estimates based on the list of ENERGY STAR certified buildings and plants (U.S. Environmental Protection Agency 2015)

[2] Authors' cumulative estimates based on the LEED project directory (U.S. Green Building Council 2015); excludes buildings certified under LEED for Homes rating systems; buildings with multiple certifications are counted as a single building.

certification requirements, such as New York City's Local Law 86, and certification policies enacted by many colleges and universities, retail chains, and other large corporations.

This increase in certifications has been motivated, in part, by the numerous purported benefits associated with green buildings and green building certification. A host of economic benefits have been observed in certified buildings, including rental and sales price premiums, reduced utility costs, lower vacancy rates, and competitive advantages linked to marketing and public image (Fuerst and McAllister 2011). Other benefits, however, have been less conclusively demonstrated. While some studies have shown a number of benefits to building occupants in certified buildings in the form of improved thermal comfort, satisfaction, and productivity (Newsham et al. 2013), others have found negligible benefit (Altomonte and Schiavon 2013). Similarly, the energy and greenhouse gas emission reductions associated with certification have been the subject of considerable scrutiny, and while some studies have shown reduced energy use in certified buildings (Newsham, Mancini, and Birt 2009), others have not (Scofield 2013). The conclusions from these studies are limited, in part, by a lack of energy performance data for a representative sample of certified and non-certified buildings (Scofield 2013).

The CBECS is a statistically representative sample of the U.S. building stock. The 2012 CBECS, which introduced questions on green building certification, could potentially be used to fill the void of energy data for certified versus non-certified buildings. The purpose of this study is to examine the validity of the survey responses to the green building certification questions in the 2012 CBECS and to explore possible factors associated with instances of false reporting.

Data and Methods

Overview of the 2012 CBECS and green building certification questions

The CBECS is the only nationally representative survey focused on energy consumption and energy-related characteristics of the U.S. commercial building stock. It is conducted by the U.S. Energy Information Administration (EIA), which is the independent statistical agency within the U.S. Department of Energy (DOE). CBECS has been conducted ten times since its inception in 1979, most recently in 2012. A statistically representative sample of buildings across the United States are selected using a multistage area probability sample, supplemented by several administrative lists of large building types such as hospitals, Federal government buildings, and airports.

The 2012 CBECS data collection resulted in 6,720 completed building cases. The survey data are collected during a computer-assisted interview with a representative from the building (the respondent) who is either the building manager or a person deemed to be "knowledgeable about the types of energy used in the building." Respondents vary in their levels of expertise, from chief building operators to store clerks. In the 2012 CBECS, well over one-third of respondents were categorized as operations, maintenance or energy managers and an additional one-third were building owners, business owners, or property managers. Questionnaire topics include building size and age, ownership and occupancy, energy sources and end uses, and types of energy-related equipment found in the building. A representation of the complete survey instrument is available online[3].

The 2012 iteration was the first CBECS to include questions about green building certification, which are labeled L14, L15 and L16 in the survey instrument. These questions were introduced following discussions with program managers for EPA ENERGY STAR and with building technology experts in DOE's Office of Energy Efficiency and Renewable Energy. The intent was to allow tracking of market penetration of certified green buildings and analysis of the characteristics and energy use of certified and non-certified buildings.

Question L14 asked: "*In the past 3 years, has this building been certified as a 'green building,' such as ENERGY STAR, LEED, or Green Globes?*" Respondent answers are shown in Table 1. The majority of respondents selected "No," and these respondents, as well as those who did not know whether their building was certified or refused to answer the question, were not asked any further questions about green building certification.

Respondents who selected "Yes" were then asked Question L15: "*Which type of certification did the building receive? Was it Energy Star, LEED, Green Globes, or did it receive some other green building recognition?*" and were directed to select all certification types that applied to their building. Respondent answers are summarized in Table 2. While the majority of respondents indicated receiving just one certification type, several respondents did select more than one type of certification. Overall, ENERGY STAR certification, either alone or in combination with another type of certification, was the most frequent certification type selected, followed by LEED.

Table 1. Respondent answers to CBECS Question L14

Survey Response				
Yes	**No**	**Refused**	**Don't Know**	**Total**
402	5,938	1	336	6,677*

*There are 6,720 total buildings in the 2012 CBECS sample, but 43 strip shopping centers were not asked the green building certification questions because they use a different estimation method than the other strip shopping centers in the sample.

[3] Form EIA-871A. www.eia.gov/consumption/commercial/data/2012/pdf/questionnaire.pdf

Table 2. Respondent answers to CBECS Question L15

Certification Type	Total
ENERGY STAR only	194
LEED only	80
Green Globes only	18
Other only	33
ENERGY STAR & LEED combined	47
All other combinations	16
Did not know certification type	14
Total	**402**

Respondents who selected "Other" were then asked the open-ended Question L16: "*What was the other type of recognition received by this building?*" Respondent answers to this question ranged widely in type and specificity, from "a green business award," to programs like BOMA 360 and ENERGY STAR Leader. One respondent indicated "Other" type of certification as a rebate for decreased electricity usage, and another cited a recycling program. Due to the inconsistent level of specificity across responses, Question L16 was not validated in this study.

Record linkage of 2012 CBECS sample with certified building directories

Survey responses in the 2012 CBECS were linked with records in publicly accessible, downloadable ENERGY STAR (U.S. Environmental Protection Agency 2015) and LEED project directories (U.S. Green Building Council 2015). While these records themselves may have errors (e.g., coverage, nonresponse, missing records, typographical errors), they are the most comprehensive listing of certified buildings available, and are therefore considered the "gold standard" for the purposes of this study. Buildings in the LEED directory outside of the scope of this analysis (e.g., buildings registered but not certified, buildings located outside of the United States, buildings certified under LEED for Homes) were filtered out prior to analysis. The Green Globes certified buildings directory was not available in downloadable form (searchable web interface only) and, as a result, Green Globes certified buildings were not able to be validated in this study (Green Building Initiative 2015).

A two-step process was developed to link the survey data and directory records. First, an approximate ("fuzzy") string matching method was used to automatically match building addresses in the 2012 CBECS with addresses in the directories. Rather than looking for an exact character-by-character match for an address, this method computes an approximate string distance between two addresses, consequently performing better in the presence of typos and nonstandard address formats than exact matching. Critically, this automated technique allows for the efficient identification of potential false negative responses from the large number of respondents in the sample who reported that their building was not certified. For this study, the edit (Levenshtein) distance between each address in the 2012 CBECS and each address in the ENERGY STAR and LEED project directories, respectively, was calculated using the **stringdist** package (van der Loo 2014) in the statistical computing software R (R Core Team 2014). The directory address with the minimum edit distance to each address in the CBECS was then selected, and compiled into a database with its corresponding CBECS address. Second, this database was then sorted by edit distance and each building in the CBECS sample was manually compared to its automatically selected directory address to determine whether these were, in fact, the same building. If manual verification determined that the CBECS building was actually in the directory, its certification year was checked to make sure it was certified within the period 2009-2012 and, if so, it was deemed "certified" for the purposes of this analysis.

Analysis methods

The buildings in the 2012 CBECS sample were cross-classified by their response to the green building certification questions and the corresponding record in the certified building directories. A generic version of the 2x2 contingency table used for this analysis is shown in Table 3. For the purposes of this analysis, each response combination to question L15 was treated as a separate dichotomous response, and survey responses of "Don't know" or "Refused" were grouped under "No."

Table 3. 2x2 table used to calculate validity indicators

	Directory Record (True Value)		
Survey Response	**Yes**	**No**	**Total**
Yes	TP	FP	TP + FP
No	FN	TN	FN + TN
Total	TP + FN	FP + TN	TP + FP + FN + TN

TP = True positives, FP = False positives; FN = False negatives, TN = True negatives

These tables were used to calculate several validity indicators. The criterion validity of a survey item is typically measured by the correlation between the respondent answers and their "gold standard" values (Groves et al. 2009, 274). Correlation coefficients range between -1.0 and 1.0 and absolute values of 0.70 or greater are generally accepted as indicative of good validity (Litwin 1995, 45). For two dichotomous variables, the correlation coefficient is known as the phi coefficient, denoted r_{ϕ}, and is used as the overall metric of validity for this study.

While the correlation coefficient provides a single overall measure of association between the survey responses and directory records, other statistical measures that can be more informative for dichotomous data were also calculated. Sensitivity, or true positive proportion, is the probability that the respondent said that the building is certified, given that the building is truly certified (TP/[TP+FN]). Specificity, or true negative proportion, is the probability that the respondent said that the building is not certified, given that the building is not truly certified (TN/[FP+TN]). Positive predictive value (PPV) is the probability that a building is truly certified, given that the respondent said it was certified (TP/[TP+FP]). Sensitivity, specificity, and PPV values closer to 100% indicate a more valid survey instrument. These statistical measures have been used previously as indicators of validity for survey responses, particularly in epidemiological surveys (e.g., Parikh-Patel et al. 2003).

Two methods were then used to examine factors that might explain the false responses. First, logistic regression was used to explore potential relationships between several predictors of interest and two binary response variables: one indicating whether the survey response is a false negative (compared to true positive), and one indicating whether the survey response is a false positive (compared to true negative). Four predictor variables describing basic information about the building and respondent were evaluated in this study: building activity, building owner type, the respondent's self-described job function, and building square footage category. Original predictor categories were collapsed in logical ways to address sampling zeros, and to facilitate more manageable results. Regressions were performed in the **survey** package (Lumley 2004) in R to account for the complex sampling design in the 2012 CBECS.

Second, CARI recordings of the interaction between the respondent and the interviewer were reviewed. CARI is a survey tool that captures digital recordings of the interaction between interviewers and respondents to validate interviewer performance and data quality. The 2012 CBECS was the first CBECS to implement CARI, and EIA selected several survey questions of interest to record, one of which was green building certification Question L14. If the respondent consented to be recorded, the CARI system captured an audio recording and a screenshot of the answer keyed by the interviewer for the selected questions. Neither the respondent nor the interviewer knew which questions were being recorded. For Question L14, a random 10 percent of the cases were recorded, yielding recordings for 19 of the 134 false positives and 6 of the 58 false negatives. While this is a small sample of the total false responses, it provides some insight into response errors that might be evident in the respondent-interviewer interaction.

Results

How valid are survey responses for the certified green building questions?

The results of the cross-classifications are provided in Table 4 and the validity indicators, along with their 95% confidence intervals, are provided in Table 5. For Question L14, 259 buildings in the CBECS sample were successfully matched to either the ENERGY STAR or LEED databases; of these, 201 were correctly identified as certified by the building respondent, for a sensitivity of 77.6%. For Question L15, the sensitivities vary by certification type, with buildings that are certified as both ENERGY STAR and LEED having the highest sensitivity (71.4%), followed by buildings certified as ENERGY STAR only (64.8%) and LEED only (54.5%). The specificities are high across both Question L14 and L15, indicating that a high proportion of buildings that are not certified were correctly identified as such. For Question L14, 335 respondents identified their building as certified, and 201 of these were successfully matched to either the ENERGY STAR or LEED databases, for a PPV of 60%. For Question L15, PPV differs by certification type, with a statistically significant difference between buildings certified under both ENERGY STAR and LEED, and buildings certified under LEED only. The correlation coefficients (r_ϕ) are only moderate, and do not exceed the commonly accepted threshold ($r_\phi > 0.70$) for the responses to be considered valid measures of the true value.

While there is no absolute threshold for “good” values of specificity and PPV, in practical terms, these results are somewhat unexpected. Given that the CBECS respondent is someone who at the time of interview is deemed to be sufficiently “knowledgeable about the types of energy used in the building” and green building certifications are generally publicized events, with many certified buildings displaying award plaques in a highly visible location, one could reasonably expect the sensitivities to be higher than these results show. The relatively low PPV values are perhaps less surprising; survey respondents exhibit a known social desirability bias, and tend to over-report socially desirable traits like positive attitudes towards environmental sustainability (Groves et al. 2009, 224).

Table 4. Cross-classification counts

Survey Question	True Positive	False Negative	False Positive	True Negative
L14				
Certified within past 3 years?	201*	58	134*	6,217
L15				
ENERGY STAR & LEED	30	12	17	6,551
ENERGY STAR only	103	56	91	6,360
LEED only	30	25	50	6,505

*402 respondents replied that their building was certified, but responses indicating certification under Green Globes or Other could not be validated, and were omitted from this analysis, leaving 335 positive responses and a total sample size of 6,610.

Table 5. Validity indicators with 95% confidence intervals

Survey Question	Sensitivity (%)	Specificity (%)	PPV (%)	r_{ϕ}
L14				
Certified within past 3 years?	77.6 (72.5, 82.7)	97.9 (97.5, 98.2)	60.0 (54.8, 65.2)	0.67
L15				
ENERGY STAR & LEED	71.4 (57.8, 85.1)	99.7 (99.6, 99.9)	63.8 (50.1, 77.6)	0.67
ENERGY STAR only	64.8 (57.4, 72.2)	98.6 (98.3, 98.9)	53.1 (46.1, 60.1)	0.58
LEED only	54.5 (41.4, 67.7)	99.2 (99.0, 99.4)	37.5 (26.9, 48.1)	0.45

What factors explain false negative and false positive responses?

Logistic Regressions. The results of the logistic regression models are presented in Table 6. The percentage of false responses, multivariate (adjusted for other variables in the table) odds ratio estimates and 95% confidence intervals are provided for each category of the four factors. The base category for each factor is listed first and indicated with an odds ratio equal to 1. Categories with odds ratios significantly greater than 1 indicate higher odds of false reporting compared to the base category; categories with odds ratios significantly less than 1 indicate lower odds of false reporting compared to the base category.

Table 6. Predictors of a false response to Question L14

	False Negative			False Positive		
Variable	**%§**	**OR‡**	**95% CI‡**	**%**	**OR**	**95% CI**
Building activity						
Office	8%	1.00		1%	1.00	
Education	40%	1.15	0.13, 10.04	5%	1.59	0.16, 15.80
Health care	14%	2.34	0.19, 28.98	1%	0.67	0.23, 1.99
Lodging	64%	59.73*	6.35, 561.48	1%	0.65	0.13, 3.30
Mercantile	75%	10173.08*	60.09, 1722301.44	1%	0.79	0.10, 6.03
Other	38%	0.21	0.01, 3.19	<1%	0.19*	0.07, 0.50
Building owner						
REIT	23%	1.00		1%	1.00	
Individual	9%	0.01	<0.01, 1.28	<1%	0.47	0.14, 1.57
Private academic	12%	0.26	0.01, 4.39	13%	12.01*	1.42, 101.47
Government	40%	15.09	0.80, 285.40	2%	2.76	0.51, 14.81
Other	28%	3.44	0.27, 43.15	1%	1.42	0.46, 4.36
Respondent job						
O&M	43%	1.00		1%	1.00	
Management	37%	0.84	0.04, 19.58	1%	0.90	0.27, 3.03
Energy management	5%	<0.01*	<0.01, 0.23	5%	1.56	0.61, 4.00
Owner/Accounting	42%	9.10	0.24, 339.19	<1%	0.34	0.01, 14.70
Other	81%	7.62	0.01, 4425.45	1%	1.69	0.03, 94.93
Square footage						
50,000 or under	41%	1.00		1%	1.00	
50,001 to 100,000	16%	0.07*	0.01, 0.69	2%	1.29	0.22, 7.61
100,001 to 200,000	44%	0.17	0.01, 3.97	6%	3.62	0.66, 19.89
200,001 to 500,000	16%	0.09	0.01, 1.13	4%	2.66	0.59, 11.92
Over 500,000	13%	0.09*	0.01, 0.97	7%	8.92*	3.64, 21.85

§Percentage of responses within each category responding falsely (survey design-adjusted)
‡Odds ratio (OR) and confidence interval (CI) estimates for multivariate regressions (survey design-adjusted)
*Significant at $p < 0.05$

In the false negative model, lodging and mercantile building activities all show significantly higher odds of a false report compared to office buildings. These building types are often part of large corporations, and a disconnect between the CBECS respondent and the individuals in the corporation responsible for managing the certification process may contribute to the higher false negative odds. Energy managers had significantly lower odds of false negative response compared to operations and maintenance (O&M) staff, and this may be indicative of a general familiarity with certification programs that could be expected among individuals in this role.

In the false positive model, the odds of a false report are higher for buildings owned by private academic institutions, compared to buildings owned by a real estate investment trust (REIT). Many academic institutions have large campuses and policies that encourage or require building certification, and CBECS respondents may have mistaken the sampled building for another, certified building on campus, or, in the case of LEED certification, the respondents could have mistaken a registered (but not yet certified) building for a certified one. Many private academic institutions also emphasize their sustainability goals to their students and staff, and the increased odds of false positive reporting could also indicate social desirability bias within these institutions. Compared to buildings in the smallest square footage category, buildings in the largest category also have significantly higher odds of a false positive response.

Computer-assisted recorded interviewing (CARI). Table 7 provides a summary of the behavior displayed by the respondent and/or interviewer in the CARI recordings for the false positive responses. The majority of respondents who provided a confident false positive response (i.e., they said “Yes” or the name of the program with no hesitation) said they were LEED certified. These could be buildings that were registered but not yet certified, designed to LEED standards but not certified, or unable to be verified as certified because they were listed in the LEED directory as “confidential,” although this study cannot confirm any of these possibilities. When the false positive response was provided hesitantly, the certification type tended to be ENERGY STAR. In some cases, the interviewer seems to encourage the hesitant respondent (e.g., the respondent says, “I think so,” and the interviewer says, “Really? Great!”). Sometimes the respondent indicates that they aren’t sure, but the interviewer still enters a positive response, suggesting social desirability bias on behalf of the interviewer, as well as the respondent. Some of the false positives seem to be driven by the perception that any activities related to building energy efficiency are the same as being certified; for example, one respondent stated that they “changed [their] ballasts and all of [their] things to go green.” The fact that this response was coded as having green certification was interviewer error. The “past three years” stipulation caused a few false positives for various reasons: because the interviewer didn’t read that part of the question, because the respondent was unsure of the time frame, or because the interviewer coded it as “yes” despite the respondent clarifying that the certification was not in the past three years.

Table 7. Respondent or interviewer behavior noted from CARI recordings of false positive cases

		Type of Certification Reported			
Respondent/Interviewer Behavior	**No.**	**ENERGY STAR**	**LEED**	**Other**	**Don’t know**
Confident response	5	0	3	1	1
Hesitant response	6	4	1	1	0
Respondent describes “green” upgrades	2	0	0	2	0
Incorrect interviewer coding	2	0	1	1	0
Poor quality recording	4	1	3	2	0

Of the six false negative recordings, two of the respondents expressed uncertainty in their negative response. One of these may be explained by the fact that it was a building on a multi-building K-12 school campus; in such cases, CBECS subsamples selected buildings, and the certification may be for a different building or for the campus as a whole. The other case may be explained by the building’s 2009 certification, which is on the threshold of the question’s “in the past three years” qualification. Three respondents answered “No” with confidence. Two of these were certified under the LEED-CI system (which certifies tenant spaces within buildings), making it possible that the CBECS respondent’s expertise was with the whole building’s energy systems and they were not aware of an individual tenant’s certification. The other confident false negative was a warehouse that is leased to tenants and the respondent was a property manager, who simply may not have been most knowledgeable about green building certification. The sixth false negative recording was a bad recording and provided no information.

Discussion and Conclusions

Three major conclusions result from this study. First, CBECS respondents, who are supposed to be knowledgeable about the types of energy used in the building, do not always know whether their building is certified. Over 20% of respondents in certified buildings did not know that their building was certified, and roughly 40% of respondents who said that their building was certified were incorrect, with variation by certification type. Second, the survey responses do not have high validity. The sensitivity, PPV, and correlation values were all in the moderate range and are not high enough for these survey questions to be considered substitute measures for records of program participation. Third, a combination of factors contributes to false response. The logistic regressions show that some variable categories are, in fact, associated with significantly different odds of false reporting compared to the base category, and the CARI recordings show that respondents had issues understanding specific aspects of the questions, and provide evidence for social desirability bias. These conclusions are limited to the green building certification variables, and cannot be extrapolated to other variables in the 2012 CBECS.

The first conclusion is a byproduct of this validation study, but is important in its own right. The benefits associated with certification have been shown to be dependent on the knowledge and behavior of building occupants and staff responsible for managing the building’s energy performance. Recent research suggests that simply knowing a building is certified can increase occupant pro-environmental behaviors (Khashe et al. 2015). The impact of falsely assuming a building is or is not certified on whole building energy performance was not evaluated in this study, and remains an area for further research.

The second and third conclusions have implications for future iterations of the CBECS. They suggest at least two possible courses of action to reduce response error. First, the wording and structure of the green building certification questions could be revised and pretested prior to the survey (these questions were not pretested prior to their introduction in the 2012 CBECS). Second, these questions could be removed from the survey and records data alone could be used to collect certification information. However, additional work is needed to develop improved record linkage techniques for the CBECS and building database records. This study represents the first use of automated record linkage with the CBECS data, and several issues were encountered that resulted in heavy reliance on manual validation. Approximate string matching can erroneously identify lexicographically “nearby” records (e.g., 123 Main Street and 124 Main Street) as matches when they are not (Gu et al. 2003). More sophisticated record linkage techniques, such as “blocking,” use additional variables to reduce false matches, but differences in the way that CBECS and certification programs define and sample buildings, combined with general response error, leads to varied reporting of likely blocking variables, like square footage and year of construction. Other issues included buildings with more than one address, such as large buildings that take up entire city blocks, and campus buildings for which the certification address is an administrative building and not the certified building itself. The relative merits of these options will be weighed in the planning stage of the 2017 CBECS, which is currently getting under way.

Critically, this study has also produced validated data on ENERGY STAR and LEED certification for use with the 2012 CBECS. While this provides information on the energy consumption and characteristics of certified and non-certified buildings that is statistically representative of the U.S. building stock, there are a number of important considerations for the use and interpretation of these data in practice. First, the survey question (and, subsequently, the validated data) asked whether a building had been certified within the past three years, but ENERGY STAR buildings need to recertify annually, and most LEED rating systems do not require recertification at all. It is not clear how these differing time horizons would impact an analysis with CBECS energy consumption data collected for the year 2012.

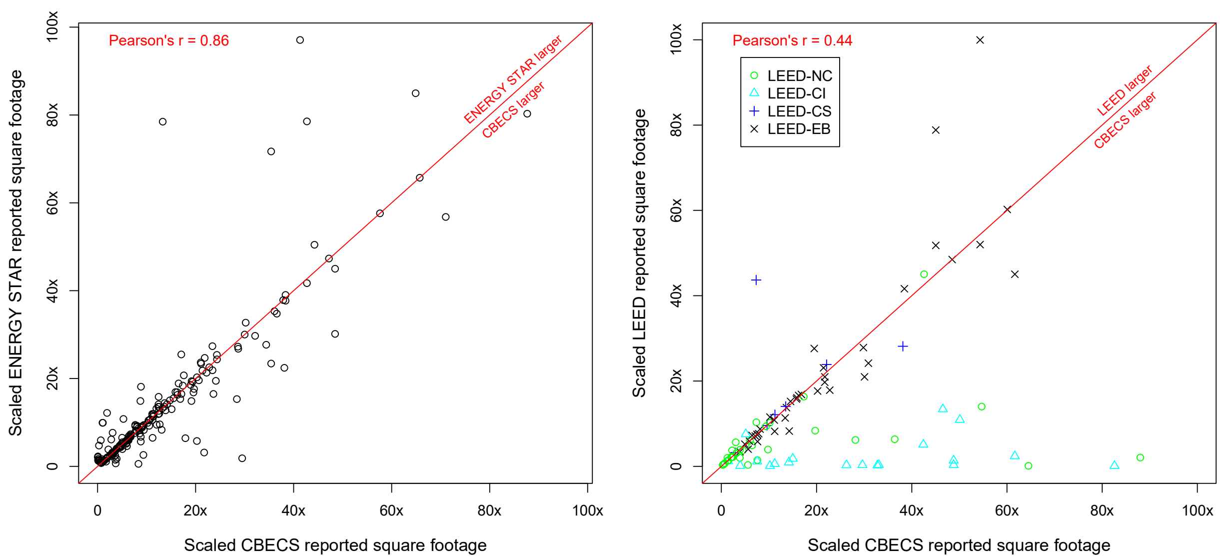

Figure 1: Building square footage as reported to CBECS (x-axis) and as reported to ENERGY STAR (left plot, y-axis) and as reported to LEED (right plot, y-axis). Building square footage in both plots has been randomly scaled to protect respondent privacy.

Second, as mentioned in the discussion on record linkage, the square footage reported to CBECS and the certified square footage do not always match up, and it is not clear how these differing spatial boundaries would impact an analysis with these data. Figure 1 plots the building square footage reported to CBECS on the x-axis (scaled by a random number to protect respondent privacy) and the building square footage (also scaled) reported to either ENERGY STAR (left plot) or LEED (right plot) on the y-axis. In the ENERGY STAR plot, the correlation (r = 0.86) is good, but there are clearly major differences in reporting for some buildings. In the LEED plot, buildings have been categorized by rating system (some LEED rating systems allow for portions of buildings to be certified). The correlation (r = 0.44) between the CBECS square footage and the LEED square footage is lower than for CBECS and ENERGY STAR, which appears to be a result of buildings certified under LEED-CI and LEED-NC with a certified square footage much smaller than the total building square footage.

Table 8. Population estimates for number of ENERGY STAR and LEED certified building in the U.S., as of and prior to 2012

Certification Program	Directory Total	CBECS Estimate	CBECS 95% CI	CBECS RSE‡
ENERGY STAR	20,468*	32,628§	15,313, 49,943	27.1
LEED	11,274†	13,745§	8,255, 19,235	20.4

*Total number of buildings in the U.S. certified in and prior to 2012, as listed in the ENERGY STAR certified building directory (U.S. Environmental Protection Agency 2015); buildings certified multiple times are counted as one building
†Total number of buildings in the U.S. certified in and prior to 2012, as listed in the LEED project directory (U.S. Green Building Council 2015); excludes buildings certified under LEED for Homes rating systems; buildings with multiple certifications counted as a single building.
§ Weighted estimates of buildings certified in or prior to 2012 based on the validated 2012 CBECS data; note that this value differs from the number of buildings certified between 2009–2012.
‡Relative standard error

A third consideration for the use of the data is the protection of respondent privacy. Despite their growing numbers, certified green buildings still make up a small proportion of the U.S. commercial building stock. Out of the 5.6 million U.S. commercial buildings estimated in the 2012 CBECS, only 0.6% had been ENERGY STAR certified and 0.2% had been LEED certified in or prior to 2012 (Table 8). Their small numbers and highly public information (e.g., certified building directories, press releases) make these buildings relatively easy to identify based on a small amount of data. Since respondent information in the CBECS is collected under the Confidential Information Protection and Statistical Efficiency Act of 2002, and any data that could potentially identify an individual building is therefore kept strictly confidential, data on green building certification will not be included in the 2012 CBECS microdata. However, information about green building certification is important in providing a complete picture of our current building stock, and the validated data may be used to construct summary statistical tables that would be made publicly available.

Acknowledgements

Amanda Webb collaborated on this project through an internship at the U.S. Energy Information Administration, completed as part of Penn State Survey Research Center's Graduate Certificate in Survey Methodology.

References

Altomonte, S., & Schiavon, S. (2013). Occupant satisfaction in LEED and non-LEED certified buildings. *Building and Environment*, *68*, 66–76.

Fuerst, F., & McAllister, P. (2011). Green Noise or Green Value? Measuring the Effects of Environmental Certification on Office Values. *Real Estate Economics*, *39*(1), 45–69.

Green Building Initiative. (2015). Certified building directory. Accessed September 22, 2015. www.thegbi.org/project-portfolio/certified-building-directory/

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2 edition). Hoboken, N.J.: Wiley.

Gu, L., Baxter, R., Vickers, D., & Rainsford, C. (2003). *Record linkage: Current practice and future directions* (CMIS Technical Report No. 03/83). CSIRO Mathematical and Information Sciences.

Khashe, S., Heydarian, A., Gerber, D., Becerik-Gerber, B., Hayes, T., & Wood, W. (2015). Influence of LEED branding on building occupants' pro-environmental behavior. *Building and Environment*, *94, Part 2*, 477–488.

Litwin, M. S. A. (1995). *How to Measure Survey Reliability and Validity. Vol. 7 [electronic resource*. Thousand Oaks: SAGE Publications.

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, *9*(1), 1–19.

Newsham, G. R., Birt, B. J., Arsenault, C., Thompson, A. J. L., Veitch, J. A., Mancini, S., … Burns, G. J. (2013). Do "green" buildings have better indoor environments? New evidence. *Building Research & Information*, *41*(4), 415–434.

Newsham, G. R., Mancini, S., & Birt, B. J. (2009). Do LEED-certified buildings save energy? Yes, but…. *Energy and Buildings*, *41*(8), 897–905.

Parikh-Patel, A., Allen, M., Wright, W. E., & Committee, the C. T. S. S. (2003). Validation of Self-reported Cancers in the California Teachers Study. *American Journal of Epidemiology*, *157*(6), 539–545.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Scofield, J. H. (2013). Efficacy of LEED-certification in reducing energy consumption and greenhouse gas emission for large New York City office buildings. *Energy and Buildings*, *67*, 517–524.

U.S. Environmental Protection Agency. (2015). ENERGY STAR Certified Buildings and Plants. Downloaded September 17, 2015. www.energystar.gov/index.cfm?fuseaction=labeled_buildings.locator

U.S. Green Building Council. (2015). LEED project directory. Downloaded September 22, 2015. www.usgbc.org/projects

Van der Loo, M. (2014). The stringdist package for approximate string matching. *The R Journal*, *6*(1), 111–122.