ENGINEER ASSESSMENT OF RESPONSE ERROR IN INTERVIEW REPORTS OF ENERGY-CONSUMPTION-RELATED CHARACTERISTICS OF NONRESIDENTIAL BUILDINGS[1]

D. French, L. Patinkin, L. Carlson
Energy Information Administration

H. Nisselson, A. Chu, and M. Berlin
Westat, Inc.

Summary

As part of a study of survey methods for measuring energy-consumption-related characteristics of nonresidential buildings, the Energy Information Administration conducted a field test in the Portland, Oregon, SMSA in the Fall of 1980. Data from the survey were evaluated by on-site inspection by engineers in a subsample of about 100 of the buildings. They examined building plans where available, made personal observations, and asked relevant questions in order to obtain answers of a quality that is involved in engineering energy evaluations.

In this paper we present some comparisons of interviewer and egineer results for that subsample. From a substantive point of view, our conclusion is that for most questions lay interviewers can provide useful information, but a few items cannot reasonably be obtained by them. There is a considerable potential for utilizing engineering data from a small subsample of a larger national survey through a double-sampling estimator; at least for improving estimates on those items for which interviewers could obtain data that are useful, but subject to larger biases than desired. Rough "upper-limit" estimates of the correlated component of response variance indicated that it was not substantial for most items.

Methodology

The sample of about 600 buildings in the original survey was selected with probabilities varying according to a measure of size based on electricity consumption (1). The assessment sample was a systematic subsample of 104 buildings from the 511 completed building interviews. Field evaluations were completed for 97 of the 104 buildings. Six engineers participated in the assessment and the work of 16 survey interviewers was involved. The assignment of interviews to both interviewers and engineers was made on the basis of convenience for the field work, and was not randomized.

A list of the data items in the assessment and the response categories for each is given in Exhibit 1. The assessment covered 36 of the 101 data items dealing with building characteristics in the original survey and 34 of the 68 data items dealing with heating, ventilation and cooling (HVAC) systems. Items such as the characteristics of building occupants for which there was no reason to expect that engineers would--on the average--obtain more accurate responses than lay interviewers were not included in the assessment.

Since the principal goal of the assessment was to measure response bias, engineers carried copies of the completed survey questionnaires with them. After they had completed their review they attempted to reconcile any discrepancies between their answers and the interviewer answers before leaving the building. The effect of reconciliation tends to result in underestimates of the response variance, and possibly also of the response bias, even if the re-interviewer is told not to look at the initial responses until after the interview questionnaire has been completed (2). The primary edit was for total square footage of the building. It was designed to approximate more accurately the error expected in a survey by duplicating the edit procedures that would be part of a well-planned survey. An effort was made to avoid arbitration changes in the interviewer or engineer responses to artificially reduce large differences between them. The edit used sketches and polaroid pictures of the buildings made by interviewers, and the reported number of floors and square footage per floor. Errors were found in both the interview and assessment data. In a few cases it was clear that the interview and assessment referred to different buildings, and these were not used in the analysis. Finally, the analysis is limited to data items for which there was a response to the item in both the survey and the assessment. As a matter of fact, however, general experience with surveys as well as this analysis suggests that it is often reasonable to interpret the absence of a response as a negative response. With this interpretation the agreement between the survey and the assessment would have been increased and the index of inconsistency would have been decreased. However, the relative bias would have remained the same.

Measures of Response Bias

Of the 45 quantitative data items (or subparts of items) examined, six items on an unweighted basis and 11 items on a weighted basis had an observed net difference twice its estimated sampling error or larger.[2] Data for items based on reports from 30 or more buildings in both the survey and assessment are shown in Table 1. Among items based on reports from 30 or more buildings, roughly half of the items had relative net biases greater than 5 percent. The items in the table all show net biases of 5 percent or more.

Of the 32 categorical data items (or subparts of items) examined, 10 items on an unweighted basis and 17 items on a weighted basis had one or more categories with a net difference twice its estimated standard error or larger. Eight items showed such differences on both an unweighted and weighted basis. The statistically significant biases concentrated among those with net biases of five percentage points or more.

For items having a relatively high (weighted) correlation a double-sampling regression estimator or difference estimator using interviewer results and engineer results for a subsample may provide consistent or unbiased estimates with considerably smaller variance than an estimate based only on the engineer subsample (3).

Measures of Response Error -- Simple Response Variance and Index of Inconsistency

The primary measure of response variability used in this analysis is the index of inconsistency, which may be interpreted as a standardized measure of response variability. The conceptual model is that of Hansen, Hurwitz and Bershad (4). This model considers each unit surveyed as having a set of potential responses which would be generated by repetition of the survey process under the same essential conditions. The assessment is treated here as if it were an independent repetition of the survey to provide estimates of the simple response variance.

The index of inconsistency, I, is defined as the ratio of the response variance to the total variance

$$I = \sigma_R^2/\sigma_x^2 = \sigma_R^2/(\sigma_R^2 + \sigma_S^2)$$

where σ_R^2 is the response variance, σ_S^2 the sampling variance, and σ_x^2 is the total variance of the characteristic. This assumes that the correlation of response deviations and sampling deviations is trivial and can be ignored.

There is not necessarily any relationship between the index of inconsistency and the bias in the survey as measured by the assessment. That is, significant biases--at the 5 percent probability level, say--may occur with a small index of inconsistency or a large one. This independence between the two measures is illustrated by the distribution of the indexes for the 45 quantitative data items in Table 2. This distribution occurs, in part, because the variance of the bias will be small or large if the index of inconsistency is small or large. The relationship is a little stronger between the estimated bias, whether statistically significant or not, and the estimated index of inconsistency. Still, the bias is relatively small (less than 5 percent) for many items with moderately large indexes. The absence of a relationship between the inconsistency of response and net bias illustrates a fairly common phenomenon in surveys; i.e., response errors often tend to be compensating and to average out over large samples.

There is a direct relationship between the index of inconsistency as estimated from the assessment and the correlation between the two responses. This relationship is illustrated in Table 3 for the 39 quantitative items with indexes under 100.

The presence of simple response variance will decrease the correlations in cross-tabulations, increase the total variance, and ordinarily will increase the value of I. Moderately large and large values of I therefore suggest that consideration should be given to reducing response variance by redefining the item, redesigning the questionnaire, improving interviewer training, or using some other approach such as double sampling.

The indexes for the categorical items are on the average higher than for the quantitative items. For the architectural items, the indexes are generally small to moderate; for items concerned with lighting systems, the indexes are modest to large; and for HVAC items the indexes are mostly moderate.

Measures of the Correlated Component of Response Variance

The index of inconsistency provides an analysis of the simple response variance to assess the reliability of responses to the survey items. If the response deviations of different units in the same trial are correlated, the impact on the total variance of survey estimates may be substantially greater than indicated by the simple response variance (4).

Valid estimates of the correlated component of response variance cannot be made from the study because the work assignments were not randomized to either interviewers or engineers, or to editors in data processing. Consequently, apparent correlations can result from differences in the work assignments among interviewers (or engineers, or editors). Nevertheless, estimates can be made that, subject to sampling variability, can be interpreted as upper limits on such correlations. An analysis for 62 data items showed significant correlated response variance at the 5 percent level for only a scattering of items, no more than would be expected by chance.

This analysis was based on a two-way

random effects model with interaction for the differences in response between the survey and assessment (5). Let d_{ijk} denote the difference for building k with the original survey conducted by interviewer i and the assessment by engineer j. Then the assumed model is

$$d_{ijk} = \alpha + \beta_i + \gamma_j + \delta_{ij} + e_{ijk}$$

where α, β_i and γ_j are, respectively, the overall mean, interviewer effect and engineer effect, δ_{ij} is the interaction, and e_{ijk} is a random error with expected value zero.

---

[1]The opinions and conclusions expressed in this presentation are solely those of the authors and should not be considered as representing the opinions or policy of any agency of the United States Government.

[2]Comparisons between the weighted and unweighted results throw some light on differences in response problems between buildings with larger measures of electricity consumption (by and large, the larger buildings) and those with smaller measures. The unweighted data are dominated by the larger electricity consumers while the reverse is true for the weighted data. For example, 28 percent of the buildings in the unweighted assessment sample have more than 50,000 square feet. The corresponding weighted estimate is 8 percent.

References

1. "Nonresidential/Commercial Buildings Energy Consumption Feasibility Study," Vol. I, Final Report, a report submitted by Westat to the Energy Information Administration, February 8, 1982.
2. Koons, D. A., "Quality Control and Measurement of Nonsampling Error in the Health Interview Survey," National Center for Health Statistics, Series 2, No. 54, 1973.
3. Cochran, William G., Sampling Techniques. Third edition, John Wiley & Sons, New York, 1977.
4. Hansen, Morris H., Hurwitz, William N., and Bershad, Max A., "Measurement Errors in Censuses and Surveys," Bulletin of the International Statistical Institute, Vol. 38, No. 2, pp. 359-374, 1960.
5. Searle, S. R., Linear Models, John Wiley & Sons, New York, 1971.

Exhibit 1. List of Variables in Engineer Assessment

Variable	Response Categories
Exposed building surfaces	Length at ground level (ft.) (sum of sides)
Height, tallest section	Feet
Number of floors	Total number
Floors below ground level	Number
Average floor to ceiling height	Feet
Total square footage	Sq. ft., or sq. ft. class interval
Exterior walls:	
Construction type	Heavy, medium, or light
Thickness	Inches
Insulation, exterior walls:	
Any in most exterior walls	Yes, No
Type	Blankets/batts, loose fill or blown, rigid plastic foam boards, foam, anything else
Thickness	Inches
Retrofit	Yes, No. If "Yes", year insulation added or type
Roof:	
Square footage	Square feet
Construction type	Concrete, wood, sheet metal
Configuration type	Peaked or sloped, flat with air space between roof and top floor ceiling, flat roof without an air space
Insulation, roof or top floor ceiling:	
Any insulation:	Yes, No
Type	Same as for exterior wall insulation
Thickness	Inches
Retrofit	Same as for exterior wall insulation
Windows:	
Percent glass, by wall/roof	Percent by building side, roof (skylights)
Percentage that can be opened	All, percentage class interval, none
Type of glass	Glaze (single, double, triple) by type of glass (clear, clear with storm windows if single glaze, tinted, reflective), something else
Percent of windows shaded	Percent by building side

Exterior lighting:	
Type	Incandescent, fluorescent, mercury vapor, high pressure sodium, low pressure sodium, metal halide, something else
Number building pays for	Number
Average wattage per fixture	Number of watts
On-off switching method	Manual, time clock, photocell, something else
Interior lighting:	
Type	Incandescent, fluorescent, something else
Average lighting level	Watts per sq. ft.
Type of control	Local manual, centralized panel(s) for building or zones in building
Interior lighting:	
Type of central control	Time clocks (automatic), manual
Conservation features	Yes, No. If "Yes", feature and month and year installed
Heating systems:	
Percent or sq. ft. heated	Sq. ft. or percent of sq. ft.
Energy conversion systems	Self-contained units, central system (boilers) within the building, central system heated outside the building, something else
Boilers:	
Number in central system	Number
Fuels used	Fuel oil, gas, coal, electricity, something else
Output	Steam under 15 psi, steam 15 psi or more, hot water under 212°F, hot water 212°F or more
Capacity	Homepower, Btu/hr, kWh, lbs. steam/hr
Humidification	Yes, No
Conservation features	Yes, No. If "Yes", features and month and year installed
Distribution/ventilation systems:	
Type of systems	Forced hot air self-contained or central units; radiant or circulated without fans by baseboards--electric, hot water or steam--radiators/convertors, heating panels; something else
Percent of heated space	Percent of heated space served by each type
Central air handling units:	
Number	Total number and by type: both heating and cooling, heating only, cooling only
Temperature control method of each type	Varying amount of air, heating air or both; mixing warm and cool air delivered in separate ducts or mixed; controlling air temperature leaving air handling units
Cooling systems:	
Percent or sq. ft. cooled	Sq. ft. or percent of sq. ft.
Types of systems	Window units, wall units, packaged units, central system with separate chillers constructed for the building, something else
Percent of cooled space	Percent of cooled space served by each type
Packaged units cool and heat	Yes, No
Heat pump utilized	All, part or none of the cooling process
Humidistat	Yes, No
Percent humidity	Relative level to which humidity controlled
Months cooled	Average number of months
Refrigeration type	Absorption, centrifugal, reciprocating, something else

Table 1. Data items with net bias in the survey compared to the assessment twice its estimated sampling error*

Data item	Unweighted basis		Weighted basis	
	Net bias[1/]	Correlation	Net bias[1/]	Correlation
Height, tallest building section	3.4 ft. (9%)	.95	4.3 ft. (18%)	.91
Percent of building exposure in glass				
Northern	6.6% (37%)	.24	5.2% (57%)	.17
Eastern			5.3% (47%)	.29
Southern			7.9% (46%)	.64
Thickness of exterior walls	-1.2 in. (14%)	.58	-0.7 in. (10%)	.63
Number of exterior lighting fixtures building pays for			7 (150%)	.39
Percent of sq. ft. air conditioned for cooling	-6.8% (12%)	.80		.84
Capacity of central refrigeration unit	-77.9 tons (13%)	.99		.97
Percent of sq. ft. heated			4.9% (5%)	.81

* Data items with 30 or more reports in both the survey and the assessment.

1/ A positive difference indicates a survey estimate higher than the assessment. Figure in parentheses shows the net bias as a percent of the assessment estimate.

Table 2. Relationship between the index of inconsistency and the net bias in the survey estimate compared to the assessment estimate

Range in index of inconsistency (Census adjective rating)	Unweighted basis		Weighted basis	
	Number of data items		Number of data items	
	Total	With significant bias	Total	With significant bias
Under 20 (small)	16	2	16	-
20-49 (moderate)	13	2	10	4
50-99 (large)	10	2	10	6
100	1	-	5	-
Over 100	5	-	4	1
All indexes	45	6	45	11

Table 3. Relationship between the index of inconsistency and the correlation between the survey and assessment responses

Range in index of inconsistency (Census adjective rating)	Number of data items	Range in correlation of survey and assessment responses	
		Unweighted data	Weighted data
Under 20 (small)	16	.82-1.00	.81-1.00
20-49 (moderate)	13	.55-.78	.58-.81
50-99 (large)	10	.17-.55	.07-.51