Weighting Adjustments Can Help With Low Response Rates, but at What Cost to Data Quality?

Chrishelle Lawrence
U.S. Energy Information Administration, 1000 Independence Avenue SW, Washington, DC 20585

Abstract
The Residential Energy Consumption Survey (RECS) was traditionally an in-person survey, but in its most recent survey cycle, in-person, web, and mail surveys were used for data collection. Planning is underway to consider a web/mail only data collection. Moving from an in-person interview to a web/mail survey for data collection poses many challenges, and one major concern is the expected decline in the response rate. To compensate for low response rates, nonresponse and post-stratification weighting adjustments are often used produce more accurate estimates. In doing this, survey weights become larger and more error is introduced into the estimates. Using results from a national pilot study that was conducted via web/mail, this research will consider the impact varying levels of total and subpopulation nonresponse have on data quality with respect to weights and estimates.

Key Words: Weighting, post-stratification, Residential Energy Consumption Survey

1. Introduction

The Residential Energy Consumption Survey (RECS) collects data on energy characteristics, usage patterns, and household demographics from a nationally representative sample of housing units. In-person interviews were the main mode of data collection until 2015, where in-person, mail, and web modes were all used. With a shift in mode, there has been great concern about the survey response rate. As with many surveys, response rates have also declined for the RECS. In 2009, the response rate was 79.0%, but the most recent survey cycle in 2015 yielded a 50.8% response rate. The next RECS, which will be fielded in 2020, will be conducted by mail and web only. To prepare for such a major change in data collection, a national pilot study was fielded to test the web and mail survey modes. The response rate for this study was 40.2%.

Weighting adjustments for nonresponse and post-stratification can help produce reliable estimates, but as response rates decline, survey weights increase to make up the difference. This research will consider the impact varying levels of total and subpopulation nonresponse have on data quality with respect to weights and estimates.

2. Using the National Pilot Study to Test Nonresponse

The national pilot (NP) study had 3,654 respondents from both single-family and multi-family homes. Single-family homes include mobile homes, single-family detached, and single-family attached homes, and multi-family homes include all apartments. From this data set, several samples were selected to measure the impact of nonresponse on the weights and estimates. Table 1 shows the data sets and their sample sizes. These data sets represent possible response scenarios, some of which have a low probability of occurrence.

For example, having all survey responses come from multi-family homes (MO) would be highly unlikely, but still possible and worth considering.

Table 1. Response Scenario Data Sets

Sample	Description	n
WO	Respondents only allowed to use the web	807
WS	Web-submitted questionnaires	2,325
SO	Respondents from single-family homes	2,870
MO	Respondents from multi-family homes	784
H_50	50 samples using half of the respondents	1,827
HM_50	50 samples using all respondents from single-family homes and half of the respondents from multi-family homes	3,262
HS_50	50 samples using all respondents from multi-family homes and half of the respondents from single-family homes	2,219

Note: Web-submitted questionnaires include the web-only respondents and those who had a choice between web and mail but chose to respond online.

Since there a multiple ways to select a dataset using half of the respondents overall or from single-family and multi-family homes, the last three response scenarios (H_50, HM_50, and HS_50) were simulated 50 times.

3. The "Re-weighting" Process

The national pilot data had a complex weighting scheme and initial design weights were adjusted for vacancy status, primary housing unit status, nonresponse, and post-stratification. Weights were then post-stratified to the following control totals from the American Community Survey (ACS): Census division, housing unit type, tenure (owner/renter), year housing unit was built, and number of bedrooms.

A simpler weighting approach was used for the samples. The initial design weights were only adjusted for nonresponse and post-stratification. The design weights were adjusted for nonresponse using the adjustment cell method, where the weights of the nonrespondents are distributed proportionally among the weights of the respondents. Address type (single-family or multi-family) and Census division were available for every housing unit on the frame and were used as adjustment cells. The weights were then post-stratified to known control totals for Census division and address type from the ACS.

Table 2 shows the minimum, mean, and maximum weights of the national pilot and the samples. When comparing the weights to the national pilot weights, as expected, the on average, all other weights were greater. The average of the samples using all single-family respondents and half of the multi-family respondents (HM_50) had the smallest average weights and the multi-family only sample (MO) had the largest average weights. The largest weight of 899,708 came from the only apartment in the Mountain North division from the sample of respondents that could only use the web (WO).

Table 2. Weight Statistics

Sample(s)	Min. Weight	Mean Weight	Max. Weight
NP	2,120	32,350	182,402
WO	47,240	146,479	899,708
WS	21,192	50,842	152,383
SO	16,922	41,188	87,806
MO	53,652	150,776	401,822
H_50	22,903	64,701	392,922
HM_50	13,523	36,238	527,527
HS_50	15,962	53,271	125,080

4. Comparing Estimates

Since single-family and multi-family homes have different energy characteristics, a sample of respondents that relies on one more heavily than the other could provide less than desirable estimates. Figure 1 shows natural gas, wood, and fuel oil or kerosene usage by sample. For each of the fuels, the multi-family only (MO) estimate was the lowest. For wood and fuel oil and kerosene use, the MO estimates were significantly lower than the NP estimates. The web only (WO) and single-family only (SO) estimates were higher than the NP estimates for each fuel.

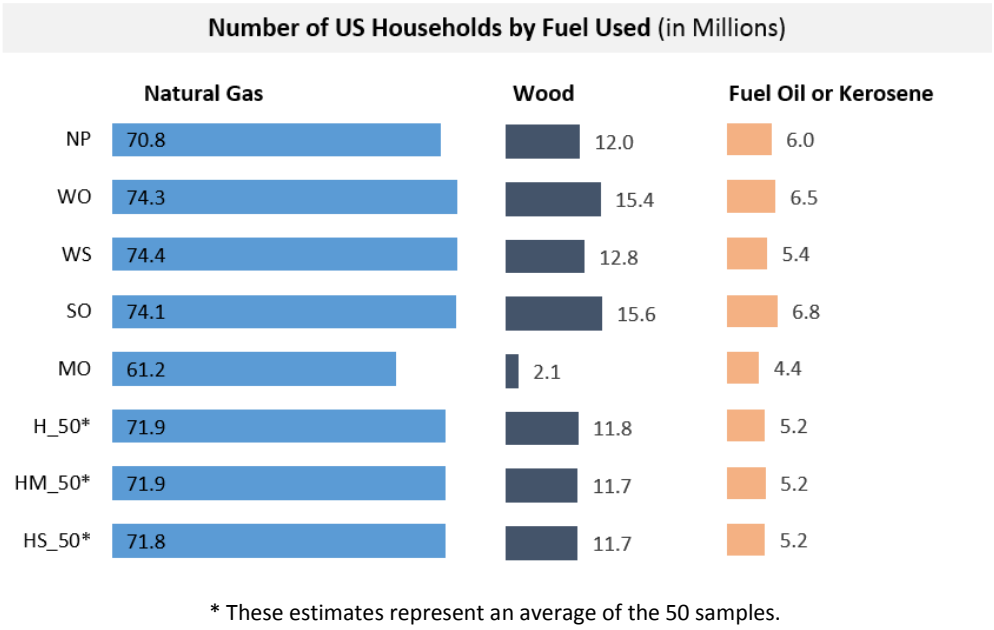

Figure 1. Natural Gas, Wood, and Fuel Oil/Kerosene Usage by Sample

Although each sample was able to produce estimates, EIA has quality standards on what can be published. Estimates are not released to the public if the number of responding cases is less than 10 or if the relative standard error (RSE) is greater than 50%. Table 3 shows the RSEs for the fuel use estimates from Figure 1. All of the RSEs fall within the publishing standards, but as with the estimates, the MO sample had the largest RSEs for natural gas and wood use. The web submitted (WS) sample had the highest RSE for fuel oil or kerosene.

Table 3. Relative Standard Errors of Fuel Use Estimates (Figure 1)

Sample(s)	Natural Gas	Wood	Fuel Oil/Kerosene
NP	1.9	5.8	7.6
WO	8.1	16.2	20.4
WS	7.5	15.4	22.5
SO	9.1	10.3	14.7
MO	10.4	37.4	21.6
H_50	3.8	8.5	13.5
HM_50	3.3	6.6	10.0
HS_50	3.6	8.6	12.0

Figure 2 shows the number of US homes owned, and Figure 3 shows the number of single-family homes in the US. Both of these estimates highlight the importance of having a diverse housing unit mix. The single-family only (SO) sample was the highest for both estimates. The multi-family only (MO) sample had significantly lower estimates, differing by over 50 million homes when compared to the NP estimates. Since the MO sample only includes housing units classified as apartments on the frame, the estimate of single-family homes should be zero. This points to some misclassification of housing type on the frame. About 16% of the housing units in the MO sample were misclassified as multi-family homes, with single-family attached homes making up the majority of these misclassified units.

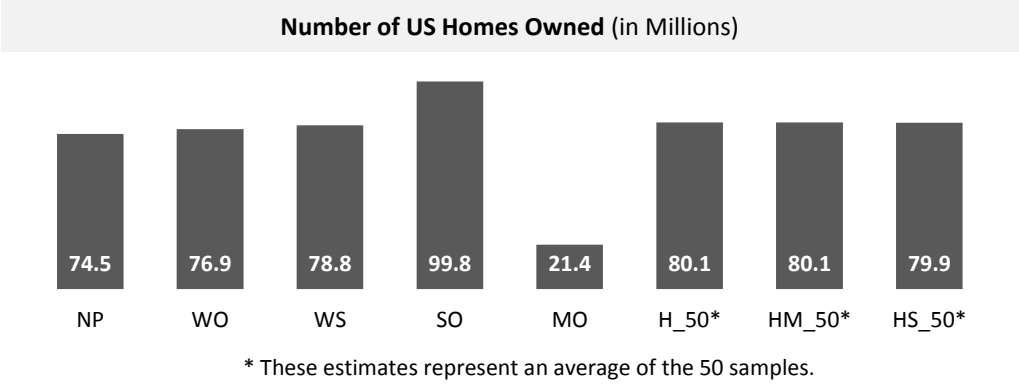

Figure 2. Home Ownership by Sample

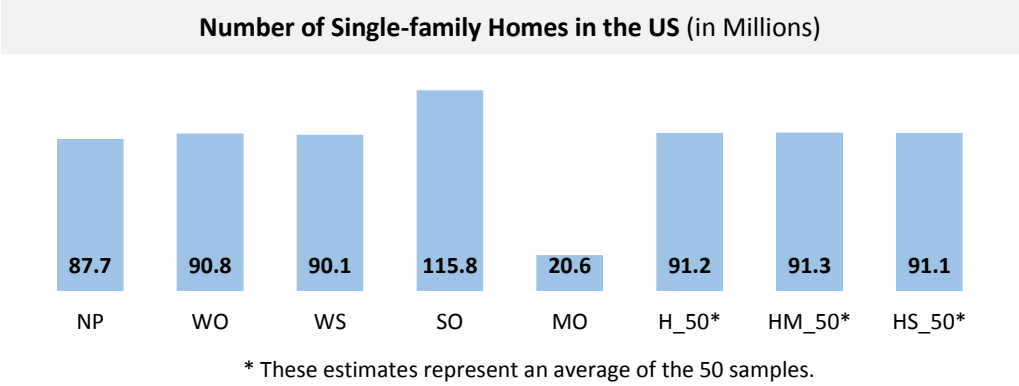

Figure 3. Single-family Homes by Sample

5. Conclusion

Varying overall and subpopulation nonresponse often yielded similar estimates within publishing and quality standards. Unreliable estimates are possible if there is not enough coverage of all housing unit types. As a result, samples of single-family only or multi-family only housing units produced estimates that were inaccurate. The national pilot study which served as the basis for this research had a small sample in comparison to the amount of respondents anticipated when a full scale RECS is conducted next year. A larger sample is expected to bring down the large weights seen in this analysis and produce reliable estimates even in the midst of declining response rates.