Weighting Mixed-Mode Data for the 2015 Residential Energy Consumption Survey (RECS)

Patrick Chen[1], Shaofen Deng[2], Neeraja Sathe[1], Lanting Dai[1], Rachel Harter[1], Phillip Kott[1]
[1]RTI International, 3020 Cornwallis Road, Research Triangle Park, NC 27709
[2]Energy Information Administration, U.S. Department of Energy, 1000 Independence Avenue, SW, Washington, DC 20585

Abstract

The 2015 Residential Energy Consumption Survey (RECS) was a stratified multistage cluster survey of housing units (HUs). RECS was designed for computer-assisted personal interviewing (CAPI) as the method of data collection. Because of difficulties experienced in the field, CAPI data collection was terminated and replaced with a web/mail data collection protocol. Nonrespondents and unfinished cases from CAPI were transferred to web/mail, and HUs in reserve replicates of sample were released to web/mail. This change imposed a challenge for weighting the combined CAPI and web/mail data. In this paper, we discuss the weighting class method to adjust for bad addresses and drop points, a latent-variable technique to predict the probability of an address corresponding to an occupied HU, and logistic regression models to estimate the probability of a HU being a primary residence. We used a calibration method to adjust for unit nonresponse and to poststratify the nonresponse-adjusted weights to the estimated number of occupied HUs from the 2015 American Community Survey (ACS) for specified HU characteristics.

Key Words: address-based sampling, mixed mode, weight calibration, weight adjustment, eligibility models

1. Introduction

The 2015 RECS was designed to measure energy consumption characteristics of U.S. households sponsored by the U.S. Energy Information Administration (EIA). The target population was occupied primary HUs in the 50 states and the District of Columbia. The 2015 RECS had a stratified three-stage sample design (EIA, 2017). At the first stage, the United States was stratified to 19 geographical domains nested within the census division. Public Use Microdata Areas were used as the primary sampling units (PSUs) (McMichael and Chen, 2015). Two hundred PSUs were allocated to the 19 geographical domains so that the target precision requirements could be met while minimizing the unequal weighting effect. The PSU samples were selected using probability proportional to size measure (PPS) method within each domain. At the second stage, four census block groups (CBGs) were selected as secondary sampling units (SSUs) from each PSU using PPS method. At the third stage, HUs were selected from CBGs using a systematic method; on average roughly eight HUs per CBG were selected. HUs were eligible if they were occupied as the primary residence of the household.

After CBG samples were selected, we obtained address-based sampling (ABS) lists and estimated the ABS net coverage rate for each selected CBG. The ABS net coverage rate was the ratio of the number of city-style addresses on the list divided by the number of HUs from the ACS. We then divided CBGs into three groups based on their ABS coverage rates:

- Group 1: ABS net coverage rates over 90% (537 CBGs)
- Group 2: ABS net coverage rates between 56% and 90% (213 CBGs)
- Group 3: ABS net coverage rates below 56% (40 CBGs)

We applied a hybrid method to construct the HU frame: for 537 CBGs in Group 1, ABS lists were used; for 213 CBGs in Group 2, ABS lists were used supplemented by a field procedure called “check housing units missed” (CHUM) (McMichael et al., 2013); for 40 CBGs in Group 3, field enumeration (FE), counting, and listing was used. CHUM had two components: check for missed units (CHUM1) and check for missed blocks (CHUM2). CHUM1 provided coverage for dwelling units missing from the ABS frame located on city blocks that had ABS coverage, while CHUM2 provided coverage for dwelling units on city blocks that had no ABS coverage. A search interval for CHUM1 began with a HU originally selected in the sample, and a search interval for CHUM2 was a block selected for this purpose. A replicate HU sample was selected and put in reserve. The goal was to complete 4,000 interviews using CAPI data collection mode.

Because of the inefficiency and difficulties of CAPI data collection, the data collection mode was switched to a web/mail protocol called Choice+ (Biemer et al., 2017). The questionnaire and a URL link were mailed to sampled HUs; they could respond using computer-assisted web interviewing or paper-and-pencil interviewing. A nonresponse follow-up was conducted for portion of web/mail nonrespondents with a shortened questionnaire.

2. Description of Mixed-Mode Data

At the time of mode change, we removed all ineligible or out of scope (OOS) cases from CAPI mode. All unknown eligible cases, nonrespondents, and pending cases were transferred to web/mail mode, except for hostile refusal cases and cases that indicated no further contact. All samples in the reserve were release to web/mail mode data collection. **Table 1** describes the mixed-mode data from three sources: CAPI only, web/mail from CAPI transferred, and web/mail from replicate samples. We completed 2,417 CAPI interviews, 957 web/mail interviews from CAPI transferred sample, and 2,312 from the replicated sample, for a total of 5,686 interviews.

Table 1. Disposition of Mixed-Mode Data for Weighting

Category	Description	CAPI Only	Web/Mail (CAPI Transferred)	Web/Mail (Replicate Sample)
Ineligible (OOS)	Ineligible Cases	N/A*	13	N/A
Completed Interviews (C)	Completed Interviews	2,417	957	2,312
Identified Not-Primary HU (S)	Not-primary HUs Identified Using Answers to Questionnaire Questions	N/A	13	94
Excluded (EXCL)	Bad Addresses and Drop Points	N/A	83	175
Nonrespondent (NR)	Nonrespondents	81	2,222	3,650
TOTAL		**2,498**	**3,288**	**6,231**

* N/A = not applicable; 736 cases identified as OOS in CAPI were not included.

3. Weighting the Mixed-Mode Data

To weight the mixed-mode data, we calculated design weights and then adjusted design weights for ineligibility, nonresponse, and coverage. Seven design weight components reflected sample selection at each stage. There were five weight adjustment factors.

3.1 Design-Based Weights

Table 2 lists all seven design-based weight components; three sources of data shared the same first five weight components. The last two CHUM-related components did not apply to the web/mail from replicated samples. We describe each component in detail below.

Table 2. Design-Based Weight Components

Design Weight Component	Description	CAPI Only	Web/Mail (CAPI Transferred)	Web/Mail (Replicate Sample)
1: PSU_WT	Inverse of PSU Selection Probability	Applied	Applied	Applied
2: SSU_WT	Inverse of Conditional SSU Selection Probability	Applied	Applied	Applied
3: SSUSUB1_FC	Inverse of FE CBG Subsampling Rate	Applied	Applied	Applied
4: SSUSUB2_FC	Inverse of FE CBG Sub-subsampling Rate	Applied	Applied	Applied
5: HU_WT	Inverse of Conditional HU Selection Probability	Applied	Applied	Applied
6: CHUM_WT	CHUM Design Weight	Applied	Applied	N/A*
7: CHUMSUB_FC	Inverse of CHUM HU Subsampling Rate	Applied	Applied	N/A

*N/A = not applicable.

1: PSU_WT This weight component reflected the first stage selection of PSUs and was calculated as the inverse of the selection probability of a PSU. The selection probability for the i^{th} PSU within the h^{th} domain was calculated as:

$$p_{hi} = n_h * \frac{s_{hi}}{\sum_{i \in h} s_{hi}}$$

where h stands for the geographical domains (h = 1, 2, …, 19); i is the index for PSUs on the frame within each domain; n_h is the number of PSUs to select in the h^{th} domain; and S_{hi} is the size measure of the i^{th} PSU, which is the estimated number of occupied HUs in the PSU based on 2013 1-year ACS data. *PSU_WT* for *hi* was calculated as the inverse of p_{hi}.

2: SSU_WT This weight component reflected the second stage of selecting CBGs from the selected PSUs and was calculated as the inverse of the conditional probability of selecting CBGs. Four CBGs were selected from each PSU. The conditional selection probability for the j^{th} CBG from i^{th} PSU within the h^{th} domain was calculated as:

$$p_{hij} = 4 * \frac{s_{hij}}{\sum_{j \in hi} s_{hij}}$$

where *j* is the index for CBGs on the frame from the *i^th^* PSU in the *h^th^* domain and s_{hij} is the size measure of the *j^th^* CBG, which is the estimated number of occupied HUs in the CBG based on 2013 5-year ACS data. *SSU_WT* for *hij* was calculated as the inverse of p_{hij}.

3: SSUSUB1_FC Among 40 FE CBGs, 28 contained at least 1,000 total HUs or covered an area that was at least 5 square miles. These 28 CBGs were subsampled or "chunked" by combining adjacent blocks to form subareas, and one subarea was randomly selected with PPS for FE. This weight component accounted for this subsegmentation. *SSUSUB1_FC* was the ratio of the total number of HUs in the CBG divided by the number of HUs in the selected subarea. For CBGs that did not require FE and FE CBGs that did not require subsampling, *SSUSUB1_FC* was set to 1. *SSUSUB1_FC* for *hij* was calculated as:

$$SSUSUB1\_FC = \frac{s_{hij}}{s_{hij_1}}$$

where s_{hij} is the number of occupied HUs in j^{th} CBG, and s_{hij_1} is the number of occupied HUs in the selected subarea based on 2013 5-year ACS data.

4: SSUSUB2_FC After the initial subsampling was completed, nine subareas still covered more than 20 square miles. These subareas were further subsampled by manually dividing them into smaller sub-subareas using the map pages created for FE. One sub-subarea was randomly selected with PPS for fielding. This weight component accounted for this further subsegmentation. For CBGs that did not require FE and FE CBGs that did not require second subsampling, *SSUSUB2_FC* was set to 1. Otherwise, *SSUSUB2_FC* for this one subarea *hij* was calculated as:

$$SSUSUB2\_FC = \frac{s_{hij_1}}{s_{hij_2}}$$

where s_{hij_1} is the number of occupied HUs in j^{th} CBG and the subarea selected in the first subsegmentation and s_{hij_2} is the number of occupied HUs in the sub-subarea selected in the second subsegmentation based on 2013 5-year ACS data.

5: HU_WT reflected the selection of a HU at the third stage from the available HU frames. The conditional selection probability for selecting a HU k from the j^{th} CBGs (or subarea or sub-subarea for FE CBGs where subsegmentation was needed) in the i^{th} PSU within the h^{th} domain was calculated as:

$$p_{hijk} = \frac{b_{hij}}{m_{hij}}$$

where b_{hij} is the number of HUs selected from the j^{th} CBG (or subarea or sub-subarea), and m_{hij} is the number of HUs or mailing addresses on the frame in the j^{th} CBG (or subarea or sub-subarea). *HU_WT* for *hijk* was calculated as inverse of p_{hijk}. For HUs added to the sample later through CHUM, HU_WT was set to 1.

6: CHUM_WT CBGs in Group 2 were subjected to the CHUM procedure for improving the coverage of the ABS frame. *CHUM_WT* was the HU-level weight component for HUs identified and added to the sample through CHUM. *CHUM_WT* was calculated differently for HUs identified and added through CHUM1 and through CHUM2. For CHUM1 HUs, *CHUM_WT* was the same as *HU_WT* of the associated HU originally selected:

$$CHUM\_WT = HU\_WT$$

For CHUM2 HUs, *CHUM_WT* was the inverse of probability of selecting blocks for CHUM2 in a CHUM CBG, calculated as:

$$CHUM\_WT = \frac{Total\ Number\ of\ Blocks\ Eligible\ for\ CHUM2\ in\ a\ CHUM\ CBG}{Number\ of\ Blocks\ Selected\ in\ a\ CHUM\ CBG}$$

This factor was applied to HUs found through the CHUM procedure. For HUs already in the sample, CHUM_WT was set to 1.

7: CHUMSUB_FC Where more than five valid HUs were identified from a single search interval of CHUM1 or CHUM2, we randomly selected five HUs and added them to the sample. *CHUMSUB_FC* accounted for this sub-selection; it was calculated separately for each search interval of CHUM1 and CHUM2 as the total number of valid HUs divided by 5. *CHUMSUB_FC* was calculated as:

$$CHUMSUB\_FC = \frac{Number\ of\ HUs\ Identified\ by\ CHUM1\ or\ CHUM2}{5}$$

For the HUs in ABS and FE CBGs, *CHUMSUB_FC* was set to 1. For originally selected HUs in CHUM segments, CHUMSUB_FC was set to 1. For the HUs added from CHUM in the CBGs where no subsampling was needed, *CHUMSUB_FC* was also set to 1.

The design-based weights (*DESIGN_WT*) for a HU in the sample were the product of the seven weight components discussed above. The design-based weights were adjusted for eligibility, nonresponse, and coverage and are discussed in **Section 3.2**.

3.2 Weight Adjustment Factors

Table 3 lists five weight adjustment factors that were used to augment the design weights to produce the final weights. Each factor is described in more detail below. Number 8 (DROPPT_FC) adjusted the insufficient addresses and drop points. Relying on mailings, web/mail survey has a well-known problem of determining eligibility of nonresponding HUs, for example vacant status and primary HU status. Number 9 (NV_FC) adjusted the weights for vacancy issue, and number 10 (PHU_FC) adjusted for not-primary residence issues. Those three adjustments were not applied to CAPI-only data. After adjusting the ineligibilities, we combined the three sources of data to perform a single nonresponse adjustment and a poststratification adjustment.

Table 3. Weight Adjustment Factors

Weight Adjustment Factor	Description	CAPI Only	Web/Mail (CAPI Transferred)	Web/Mail (Replicate Sample)
8: DROPPT_FC	Insufficient Address and Drop Point Adjustment Factor	N/A	Applied	
9: NV_FC	Not-Vacant Adjustment Factor	N/A	Applied	
10: PHU_FC	Not Primary HU Adjustment Factor	N/A	Applied	
11: NR_FC	Nonresponse Adjustment Factor	Applied		
12: PS_FC	Poststratification Adjustment Factor	Applied		

8: DROPPT_FC Eighty-three HUs transferred from CAPI, and 175 HUs from the replicate sample were drop points or had insufficient addresses for mailing. Those HUs were excluded from the web/mail data collection because they could not be contacted by mail. To adjust the weights for the cases excluded as drop points or insufficient address information, we applied a ratio adjustment method at the domain level. *DESIGN_WT* was the input weights. *DROPPT_FC* for *ijk* in domain *h* was calculated as:

$$\begin{cases} = \dfrac{\sum_{ijk \in C \cup S \cup NR \cup EXCL} DESIGN\_WT}{\sum_{ijk \in C \cup S \cup NR} DESIGN\_WT}, if\ ijk \in (C \cup NR \cup S) \cap h \\ \\ = 0, if\ ijk\ \in (EXCL \cup OOS) \cup h \end{cases}$$

In calculating this adjustment factor, we assumed that dropped "bad address" cases were eligible in the same proportion as the rest of the sample. Please refer to **Table 1** for the definitions of *OOS*, *C*, *EXCL*, etc. The bad address and drop point adjusted weights (*WEBMAIL_WT*) were the product of *DESIGN_WT* and *DROPPT_FC*.

9: NV_FC The vacant status for most web/mail nonrespondents was unknown. There are two sources we could use to determine the vacant status for web/mail nonrespondents. ABS frame included a vacancy status indicator, but it was not accurate enough for our purposes. It was even worse in areas with higher occupancy turnover rate. Another source was the USPS undeliverable notice. However, this information was inconsistent and incomplete. We used a latent class model (Biemer et al., 2016) and combined the information from the ABS frame indicator, USPS undeliverable notice, and HU response indicator to estimate the not-vacant probability. The estimated not-vacant probability was the adjustment factor to adjust unknown vacancy, *NV_FC*.

10: PHU_FC Similar to the vacancy status, the primary HU status for most web/mail nonrespondents was unknown. We defined a 0/1 primary HU indicator and used it as the dependent variable to fit a logistic regression model with some HU- and CBG-level characteristics as independent variables. The predicted probability of primary HU was applied to the HUs with unknown primary HU status as the adjustment factor, *PHU_FC*.

The not-vacant and not-primary HU adjusted weights ($ELIG\_WT$) were the product of *WEBMAIL_WT*, *NV_FC*, and *PHU_FC*; they were the input weights for the nonresponse adjustment.

11: NR_FC Failure to obtain the HU interview from eligible HUs was known as unit nonresponse. To reduce the risk of nonresponse bias, a nonresponse adjustment was implemented to adjust the respondent weights to the weighted distributions over various characteristics based on all eligible HUs. We applied the generalized exponential model (GEM) (Folsom and Singh, 2000; Kott and Liao 2012) calibration method to perform the nonresponse adjustment. GEM is a generalization of logit method (Deville and Sàrndal, 1992) and constrained exponential model method (Folsom and Witt, 1994). GEM allows the user to put bounds on the adjustment factor. Moreover, the bounds can be unit specific, the user can have control over extreme weights during the nonresponse or poststratification adjustment in a way that overcomes the limitations of traditional trimming and smoothing. *NR_FC* in GEM was calculated as:

$$\begin{cases} = \dfrac{L + \exp\left(\mathbf{g}^T \mathbf{z}_k\right)}{1 + \dfrac{\exp\left(\mathbf{g}^T \mathbf{z}_k\right)}{U}}, if\ hijk \in C \\ \quad = 0, if\ hijk \in NR \end{cases}$$

where *L* and *U* are specified bounds for the adjustment factor; $\mathbf{z}_k$ is a vector of predictor variables including an intercept (or the equivalent); and **g** estimates the model parameters associated with those variables.

We identified candidate variables for GEM first. We fit logistic regression models where the dependent variable was the response indicator and the candidate predictors were the explanatory variables. We selected explanatory variables that were significant predictors of response propensity as the initial variables for GEM. The variables kept in the nonresponse adjustment model were:

- RECS geographical domains
- PSU level proportion of detached single family HU (SFHU): <24.05%; >=24.05%
- PSU level proportion of HUs using natural gas as the major heating fuel (GAS): <73.80%; >=73.80%
- CBG level dominate HU year built: Before 1950; 1950-1969; 1970-1989; 1990 or after
- Data source: CAPI sample; replicate sample
- HU level HU type: Single; multiple
- Interactions
  - Census Region by GAS
  - Census Region by SFHU

The nonresponse adjusted weights (*NR_WT*) were the product of *ELIG_WT* and *NR_FC*; they were the input weights for the next poststratification adjustment.

12: PS_FC The nonresponse adjusted weights (*NR_WT*) were poststratified to the estimated number of occupied HUs derived from the 2015 ACS. We applied GEM to calculate the poststratification adjustment factor (*PS_FC*) as:

$$PS\_FC = \frac{L + \exp\left(\mathbf{g}^T \mathbf{z}_k\right)}{1 + \dfrac{\exp\left(\mathbf{g}^T \mathbf{z}_k\right)}{U}}$$

where *L* and *U* are specified bounds for the adjustment factor; $\mathbf{z}_k$ is a vector of predictor variables including an intercept (or the equivalent); and **g** estimates the model parameters associated with those variables. In the poststratification model, we had the following variables:

- RECS geographical domains
- HU type: Single detached; single attached; multiple units with 2-4 units; multiple units with 5+ units; mobile home
- Ownership: Own; rent
- Number of bedrooms: 0 or 1; 2; 3; 4; 5+
- HU year built range: Before 1950; 1950-1959; 1960-1969; 1970-1979; 1980-1989; 1990-1999; 2000-2009; 2010 and after

- Major heating fuel type: Natural gas; electricity; fuel oil; wood; propane; other fuel; no fuel used
- Interactions
    - Census division by HU type
    - Census division by ownership
    - Census division by number of bedrooms
    - Census division by HU year built range
    - Census division by major heating fuel type (collapsed to four levels[1]: Natural gas; electricity; all other fuel; no fuel used)

4. Results and Discussion

The fully adjusted final analysis weights for 2015 RECS mixed-mode data were the product of all 12 weight components. **Table 4** displays the weight distribution of the final analysis weights; the final analysis weights had an unequal weighting effect of 1.30. Because the not-vacant and primary HU adjustment factors were relatively small, those two adjustment steps could have been skipped and the final poststratification adjustment would have taken care of any imbalance. Another thing we could do differently is that we could have two separate nonresponse adjustments, one for CAPI only and web/mail from CAPI transferred, and the other for web/mail from replicate sample. Most CAPI transferred cases had already been worked during CAPI, but replicate samples were fresh samples and might have different nonresponse patterns.

For estimating variances of survey outcomes, we also developed 96 balance repeated replication weights.

Table 4. Distribution of Final Analysis Weights

Statistics	Final Analysis Weights
Max	158,079
95%	41,732
75%	25,276
50%	18,141
25%	13,292
5%	8,380
Min	984
Mean	20,789
n	5,686
SUM	118,208,250
UWE	1.30

Acknowledgment

This work was funded by EIA, Department of Energy, under Contract No. DE-EI-0000515.

[1] No sample for the "No Heating" category in New England, Mid-Atlantic, East North Central, West North Central, East South Central, North Mountain, and South Mountain. "No Heating" categories for South Atlantic and West South Central were collapsed.

Disclaimer

The views expressed in this document do not necessarily reflect the official policies of EIA or Department of Energy, nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

References

Biemer, P., Murphy, J., and Kott, P. (2016). Estimating mail or web survey eligibility for undeliverable addresses: A latent class analysis approach. In *JSM Proceedings* (pp. 1166-1172). American Statistical Association.

Biemer, P.P., Murphy, J., Zimmer, S., Berry, C., Deng, G, and Lewis, K. (2017). Using bonus monetary incentives to encourage web response in mixed-model household surveys. *Journal of Survey Statistics and Methodology*, Oxford Academic. (2017) 0: 1-22.

Deville, J. C., and Särndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87*(418), 376-382.

Folsom R. E., and Singh A. C. (2000). A generalized exponential model for a unified approach to sampling weight calibration for outlier weight treatment, nonresponse adjustment, and poststratification. In *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 598–603.

Folsom, R. E., and Witt, M. B. (1994). Testing a new attrition nonresponse adjustment method for SIPP. In *Proceedings of the 1994 Joint Statistical Meetings, American Statistical Association, Social Statistics Section,* Toronto, Ontario, Canada (pp. 428-433). Alexandria, VA: American Statistical Association.

Kott, P., and Liao, D. (2012). Providing double protection for unit nonresponse with a nonlinear calibration-weighting routine. *Survey Research Methods, 6*, 105-111.

McMichael, J., and Chen, P. (2015). Using census public use microdata areas (PUMAs) as primary sampling units in area probability household surveys. In *JSM Proceedings, Survey Research Methods Section* (pp. 2281–2288). Alexandria: American Statistical Association.

McMichael, J. P., Shook-Sa, B. E., Ridenhour, J. L., and Harter, R. (2013). The CHUM: A frame supplementation procedure for address-based sampling. In *2013 Research Conference Papers* Washington, DC. Available at http://www.fcsm.gov/events/papers2013.html.

U.S. Energy Information Administration (EIA). (May 2017). 2015 RECS Technical Documentation Summary. Available at https://www.eia.gov/consumption/residential/reports/2015/methodology/pdf/methodology_report.pdf