ADDRESSING NON-RESPONSE BIAS: EXPERIENCES FROM THE CALIFORNIA RASS

Tami Rasmussen, KEMA
Wendy Tobiasson, KEMA

ABSTRACT

Often researchers do not address the potential for non-response bias or they ignore it due to the perception that methods to account or correct for it are cost-prohibitive. This paper will describe a study where a multi-modal non-response survey effort was successful in minimizing the bias associated with a single-mode survey approach.

In 2003, California's investor-owned utilities and the Los Angeles Water and Power Company jointly conducted a residential appliance saturation survey. The objective of the study was to inventory residential equipment and usage patterns as well as to model overall energy use by appliance. To maximize study effectiveness, several methods were employed to collect the large volume of data used in this study. The primary survey was a direct-mail survey that yielded 18,970 respondents. To correct for potential non-response bias, the study included a follow-up effort, capturing results for 2,183 non-responders to the primary survey. The follow-up effort was delivered using two different techniques: an escalating series of mail, phone, and in-person surveys for customers in areas that could be clustered together and an incentivized priority mail package with follow-up phone surveys for customers in less populated and more remote areas.

From reviewing the final survey results, it is clear that the non-response captured a somewhat different component of the market than was captured with the direct-mail approach. In general, non-respondents had energy usage and major equipment holdings similar to direct-mail participants but differed significantly in that they were less likely to be property owners, less likely to be using energy-efficient lighting, more likely to be non-English speaking, more likely to be ethnically diverse, and less educated overall. It follows from this that the direct-mail campaign was most successful with individuals who were more aware of energy efficiency, were more motivated because of their ownership, more educated, and more capable of handling an English-language survey. These results suggested that non-respondents were in fact different from the initial pool of respondents and that follow-up efforts were justified in order to produce results representative of the population.

This paper explores the costs and benefits associated with non-response follow-up methods, providing researchers and policymakers with lessons learned about the various approaches. This paper will encourage researchers to explore the tradeoffs of directing resources towards non-response follow-up and will provide them with lessons learned on various approaches that may be used to contact non-responders.

INTRODUCTION

The California residential appliance saturation survey (RASS) study was designed to address non-response bias using a multi-tiered data collection effort. This proved successful as the responses differed between the two groups. Because the RASS was a large-scale study targeted at a wide set of customers, the results help convey the importance of including non-response follow-up efforts in similar surveys. Study design must strike a balance between cost effectively planning to capture all areas of the target

population, focusing survey design efforts to maximize research benefit while minimizing customer impacts, and minimizing non-response bias using tiered data collection methods.

Study Methodology

In an effort to consolidate planning across California, the California Energy Commission (Energy Commission) sponsored a statewide RASS. While the study was overseen by the Energy Commission, there were five utility sponsors including: Pacific Gas and Electric (PG&E), San Diego Gas and Electric (SDG&E), Southern California Edison (SCE), Southern California Gas Company (SoCalGas), and Los Angeles Department of Water and Power (LADWP).[a] These five sponsoring utilities serve energy to over 10 million California households, which is 88% of the total households in the state according to the 2000 Census. [1]

Using a statewide survey instrument provided the Energy Commission and other parties with a consistent set of questions and study results to use for statewide planning and cross-utility comparisons. In addition, the sample includes sufficient data to enable utility-specific analyses.

The project required a cooperative effort among the sponsors to create a unified research plan, program materials, and implementation strategy. The sponsors shared project costs and final results. Each utility provided the data necessary to create a unified sampling plan and provided specific information for customers who were selected for the sample. To ensure individual customer anonymity, study participants were assigned a generic identification number that included details about their sampling strata. Respondent ZIP codes are the only other information that is generally available in the final study database as to the customer's location.

The project used a hybrid data collection strategy. Most of the survey data were collected using a mail survey. This consisted of mailing survey packages to approximately 100,000 customers with an individual electric meter and following up with a second survey package to those who did not reply to the first mailing. Telephone interviewing was used to gather data initially from electrically master-metered accounts (those with a single meter serving multiple dwellings or units), who were then mailed a survey package.

To reduce the non-response bias that was likely to occur from a mail survey alone, a second step of surveying efforts was pursued on a sample of 5,000 non-respondents to the mail survey. We ultimately surveyed a total of 2,183 of these non-respondents using either a third mail survey with an incentive, a telephone interview, or an in-person interview at the home.

Sample Design for Individually Metered Accounts

The study employed a stratified modified proportional sampling design using the utility population data from the four sponsoring electric utilities.[b] The purpose of this was two-fold: to ensure smaller market segments are assigned a sufficient sample to maintain reasonable precision levels and to lower the overall sampling error by using known information for the various market segments to reduce the number of sample points required to achieve precision goals. While Neyman Allocation is often used to assign sample points across consumption strata, KEMA opted for a modified proportional allocation method with minimum sample sizes by segment to better meet the multiple needs of the RASS data.

---

[a] Sacramento Municipal Utilities District was also invited to participate, but declined.

[b] SoCalGas' population data was not part of the electrically based sampling plan.

Separate strategies were used for the individual and master-metered sample frames. For individually metered customers, there were 105 strata, based on 6 variables: electric utility, age of home, presence of electric heat, home type, dwelling type combined with usage, and Energy Commission forecast climate zone.

By assigning a minimum number of target completes per sampling cell, the sample ensured representation in each of the population segments. Likewise, higher mail-out rates were specified for groups that were likely to have lower response rates based on experience from prior RASS studies. The total individually metered sample was set at 100,999. The individually metered customers' response rate was 18.8% overall, with 2,260 responses coming from the non-response portion, which is described later. While the response was lower than projected, the overall volume of responses yields results with reasonable confidence bounds. Sampling variability at the 90% confidence for the worst case (proportional estimate of 50%) varies across utility as shown in Table 1.

Table 1. 90% Confidence Bounds by Utility for Individually Metered Households (Includes Non-Response Follow-up Results)

Electric Utility	Population	Percentage of Total Population	Actual Completes	Percentage of Total Completes	90% Confidence Bounds (+/-)
PG&E	4,047,694	41%	9265	44%	1.9%
SCE	3,857,361	39%	7979	38%	2.0%
SDG&E	1,128,806	11%	2527	12%	3.7%
LADWP	879,001	9%	1382	7%	4.5%
Total	9,912,862	100%	21153	100%	1.2%

Sample Design for Master-Metered Accounts

Master-metered accounts were surveyed differently, depending on the type of units the account serves. All master-metered accounts were assigned to a stratum based on a proportional sample design that approximates the ratio of target completes to the number of units or dwellings (not accounts). For this study, we stratified master-metered accounts by utility and by type of account: master-metered accounts serving 2 to 4 units, mobile home parks with 5 units or more, multi-family complexes with 5 to 20 units, and multi-family complexes with more than 20 units.

Accounts serving two to four units were surveyed similarly to individually metered accounts. Master-metered accounts serving more than four dwelling units were surveyed using a two-stage method. In the first stage, we conducted telephone surveys with a facility manager of the multi-family complex or mobile home park to obtain data on the common area equipment and to obtain mail addresses for the dwelling units served by the account.

The second stage involved selecting a sample of units. We collected information on occupants from the facility manager and assigned the number of units per account type based on expected response rates. In all master-meter cases, we randomly selected addresses within the complex, entered information provided by the facility managers, and mailed the surveys. Customers completed the remaining portion of the unit-specific questions. The overall response rate for the master-metered sector was 13.7%.

Table 2 presents the sampling variability at the 90% confidence for the worst case for the master-meter accounts at the utility level.

Table 2. 90% Confidence Bounds by Utility for Master-Metered Homes

Electric Utility	Master Meter Population	Percentage of Total Population	Actual Completes	Percentage of Total Completes	90% Confidence Bounds (+/-)
PG&E	203,394	47%	382	50%	5.7%
SCE	153,954	35%	261	34%	6.0%
SDG&E	61,400	14%	120	16%	12.4%
LADWP	16,198	4%	4	1%	na
Total	434,946	100%	767	100%	4.0%

Direct Mail Survey Implementation

Survey packages were mailed out to all targeted customers. The survey package consisted of a cover letter, business reply envelope, scannable survey, and an outer envelope. There was no incentive. The survey was designed to capture as much information as possible to produce saturation estimates and calculate unit energy consumption (UEC) estimates as well as to collect some market intelligence information to help utilities better understand customers. The Energy Commission and all utilities participated in the design of the survey package.

Survey packages were mailed out third class. A second batch of survey packages was mailed to customers who had not responded several weeks after the initial mailing. The second batch of surveys was also sent via third-class mail with the same materials. All surveys were processed with optical scanning equipment to expedite processing and ensure consistency.

Non-response Follow-Up

To reduce the non-response bias that was likely to occur from the mail survey, a second surveying effort was made on a sample of 5,000 non-respondents to the mail survey. These customers were targeted using a combination of mail, telephone, and/or in-person contacts.

To gather non-respondent data from across the state most cost effectively, the non-response sample was divided into two groups. More densely clustered ZIP codes were sampled with between 10 and 20 households per ZIP code, for a total sample of 4,395 in 229 ZIP codes. Less densely clustered ZIP codes were sampled with between 1 and 10 households per ZIP code, for a total sample of 605 in 236 ZIP codes. Clusters were selected with probability proportional to size such that each non-respondent had an equal chance of being selected for the follow-up sample.

Clustered households were targeted using a third questionnaire mailing with a $1 incentive, follow-up phone calls, in-person solicitations, and door hangers with survey information in an escalating process to obtain a target number of responses within each sampled area. Data collection for non-clustered households included a USPS Priority Mail package sent to each household with a $5 incentive and the promise of a $15 incentive upon receipt of the completed questionnaire. The non-clustered households were then targeted with follow-up telephone calls but did not receive in-person visits because of their disparate locations.

Response to the non-response follow-up effort was 47% of the eligible customers in the non-response sample. A total of 2,260 non-response surveys were completed. Almost a third of these responses came from the mail. The non-response first-class package yielded a 10.6% response. The priority mail package with high incentives yielded an impressive 32.4%. Response to telephone interview efforts was

12%, which was somewhat lower than expected because of difficulties getting customer telephone numbers. In-person interviews yielded a response of approximately 34%.

Effect of Combining the Main Sample and Non-response Follow-Up Sample

To combine the results from the initial sample and the non-response follow-up efforts, the study combined the weights from both components to create a set of individual weights that represents the number of households that each participant represents. Instead of fully weighting the non-respondent results to represent all non-respondents, the follow-up sample weights were reduced in a systematic approach. This assumed that the follow-up sample represents only those customers who would respond to the follow-up survey but not to the main survey, rather than assuming the follow-up respondents represent all non-respondents to the main survey. This approach improved overall precision and reduced the likelihood of individual outlier cases in the non-respondent sample from skewing overall results. While more complex strategies for combining the samples were considered, this simplified approach maximized precision while avoiding the reliance on tertiary results (i.e., 2000 Census data) to re-weight the sample. This decision was made after comparing results to the census and determining that the RASS findings reasonably followed established trends. Figure 1 in the Census comparison section provides the comparison results with the Census information.

The non-response follow-up proved to be a successful way to capture a segment of the population underserved by the direct-mail campaign. Table 3 shows several key results for customers by dwelling type and survey method.

In general, non-respondents had similar energy usage and major equipment holdings as direct-mail participants but differed significantly in that they were less likely to be property owners, less likely to be using energy-efficient lighting, more likely to be non-English speaking, more likely to be ethnically diverse, and less educated overall. It follows from this that the direct-mail campaign was most successful with individuals who were more aware of energy efficiency, were more motivated because of their ownership, more educated, and more capable of handling an English survey. The non-response follow-up was able to get to more Spanish-speaking customers. While the non-response follow-up adds significant cost to a project of this magnitude, the fact that customers differ in these ways indicates that it is a wise step to take to minimize non-response bias found in a single-method survey approach.

COMPARISON TO CENSUS DATA

To understand how the results correspond to the population of California, we compared 2000 Census data to the RASS results. Overall, the comparison of the RASS demographic information to the 2000 Census data is reasonable, and the sampling plan yielded a set of customer respondents that closely mirror the population at large. The most notable area where the study appears to fall short is in the single-occupant rental market. The shortfalls occur predominantly in the young-adult age groups. Because the results aligned with census data, the study group decided to keep the initial sample weights and not post-stratify the results.

A series of figures is included as Figure 1, which compare demographic results from the Census to the complete RASS survey sample versus the RASS mail survey only. As shown, the follow-up sample helped to increase the percentage of renters, Hispanic households, and households with 1 to 2 persons in the overall RASS sample, bringing the overall demographic profile closer to the Census.

Table 3. Comparison of Results by Surveying Method and Dwelling Type

	Single Family		Multi-Family (2-4 Units)		Multi-Family (5+ Units)		Mobile Homes	
	Initial Mail	Non-Response	Initial Mail	Non-Response	Initial Mail	Non-Response	Initial Mail	Non-Response
Completed Surveys	12,599	1,225	2,979	409	2,866	512	526	37
Weighted to Population	2,363,823	3,693,704	524,317	1,155,001	513,069	1,463,655	95,691	103,602
Average Electric Consumption	7,248	7,160	4,429	4,201	3,689	3,969	6,271	6,531
Average Gas Consumption	547	538	341	338	215	216	491	478
Average Dwelling Size	1,837	1,755	1,156	1,061	925	914	1,258	1,083
Average Dwelling Age	14.5	18.9	24.0	24.8	28.4	34.6	19.4	27.9
Average Number of People	2.88	3.42	2.53	2.74	2.10	2.68	2.30	2.22
Average Number of Seniors	0.53	0.30	0.38	0.13	0.37	0.15	0.74	0.42
Average Income	73,389	68,714	54,246	47,346	45,388	41,702	30,971	28,807
Owners	91%	81%	50%	26%	26%	13%	87%	89%
Central Cooling	50%	47%	40%	33%	41%	31%	60%	38%
Gas Space Heating	85%	89%	77%	75%	46%	54%	57%	56%
All Exterior Walls Insulated	56%	61%	45%	48%	43%	44%	65%	59%
CFL Penetration	63%	50%	55%	42%	51%	37%	57%	51%
Primary Language English	92%	80%	85%	67%	87%	69%	95%	81%
Head of Household Hispanic	12%	26%	17%	36%	13%	33%	9%	20%
College Grad or Higher	53%	44%	47%	39%	50%	36%	23%	18%

COST-EFFECTIVE STUDY PLANNING

While there are numerous ways to collect RASS or other large project data, it is important to consider the most cost-effective way to capture the necessary information. RASS surveys that target utility populations are most commonly done using mail surveys. Mail is advantageous because it offers a low-cost data collection approach and is very scalable at minimal incremental cost. Mail surveys can be sent repeatedly, can have a standard layout, and can be completed at the convenience of the respondent. Phone surveys are cost advantageous when processing volumes are relatively low (1,000 total completed surveys or less) and are preferable when complex questions are desired in certain cases (for example, to probe more information on program participation for those who have participated), or when the results need to be collected very quickly. While on-site surveys are the most expensive, they offer the best assurance of accurate information and provide opportunities to measure or otherwise actually record information from the site. Figure 2 demonstrates a sample comparison of base costs to perform a RASS type survey implementation based on a survey that would take about 20 minutes to complete. Because there are numerous factors affecting costs, this table is primarily meant to show the volumetric affects on cost as well as the relative cost of each method. These costs are for gathering and data entering the survey and do not include management or analysis elements of the project.

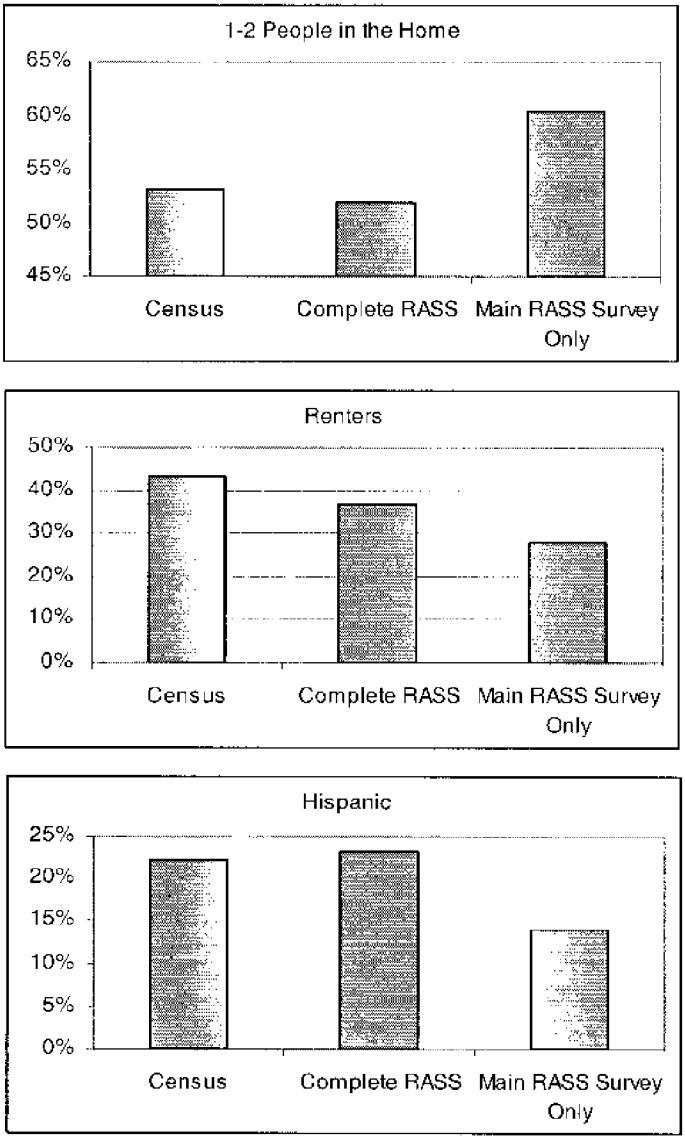

Figure 1. Comparison of RASS Results to 2000 Census Results

Other important considerations in the design of the study include understanding the likely response from various market segments. For example, single-family owners and senior citizens respond at much higher rates than renters in multi-family dwellings. If information is available at the population level, the sample may be designed to account for these expectations by oversampling segments that are expected to respond at lower rates. Another example is that non-English speakers tend to respond to mail surveys at a lower rate even when they are targeted in their native language. While choosing a single survey method (mail, phone, or on-site) simplifies the study implementation, a multi-pronged approach provides the advantage of capturing a wider audience by targeting with a range of methods that in turn helps to minimize inherent response biases. In the California RASS, 34% of initial mail respondents had

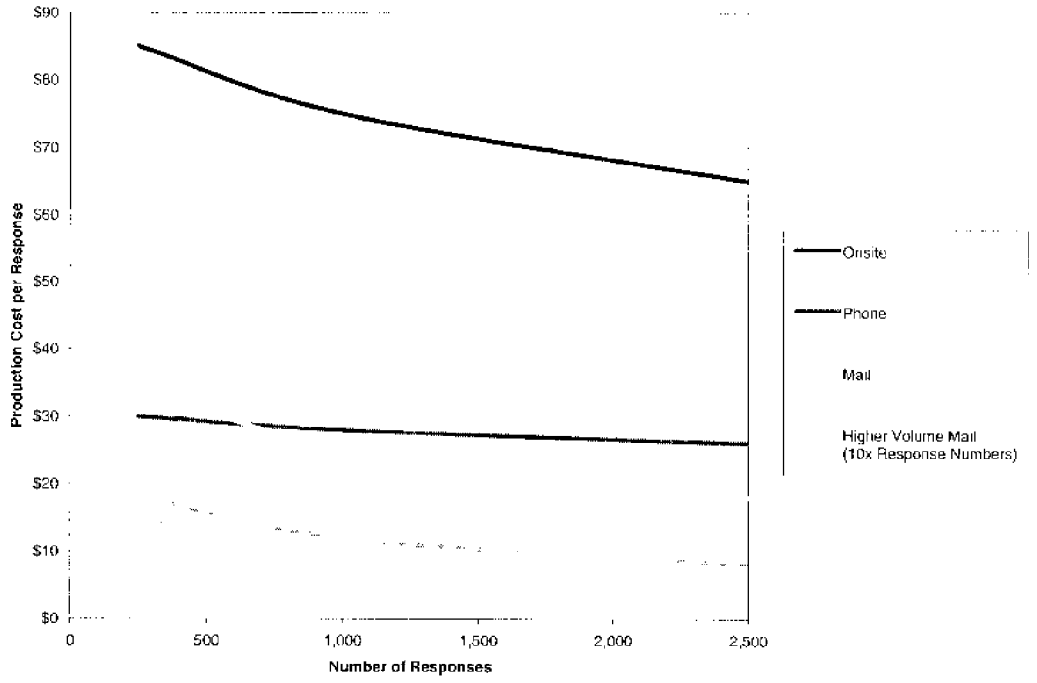

Figure 2. Sample Production Costs for a Range of Survey Methods

at least one senior in the home, compared to 25% of all respondents (including the non-response sample), illustrating how the multi-modal follow-up effort helped to correct for the higher response from seniors in the main sample.

Whatever single or combination method of surveying is chosen, the effect of non-response bias will likely be significant. A non-response follow-up survey can be implemented to address this bias. Because non-response efforts are usually much more expensive on a per-unit basis (on the order of 10:1), study designers must consider the tradeoffs between increased accuracy from directing resources towards a follow-up sample and volume of responses from the main survey. If the main survey approach is expected to achieve upwards of a 50% response rate, the bias may be small. Results from the main survey can be compared to Census data or other population statistics to provide evidence of the likely effect of the bias. If results are found to be very similar to population-level statistics, non-response follow-up may not be necessary. Where the main survey achieves much lower than a 50% response rate, the bias associated with the non-response population may be large, and study designers should consider conducting follow-up surveys with non-responders to the main survey.

The non-response survey sample must ultimately be large enough to (1) allow for meaningful comparisons between the main and follow-up surveys, and (2) be able to influence the overall estimates when combining the main and follow-up samples. One could conduct a formal optimization to determine the appropriate follow-up sample size, but the extent of the bias must be known. Since the size of the bias is typically unknown, a rule of thumb that may be used is to set the follow-up sample size at between 5 and 10% of the main survey sample, with the percentage being higher when the main sample size is smaller. For a comprehensive explanation of double sampling, including optimum allocation and comparison with single sampling, the reader is referred to Cochran's Sampling Techniques

For the California RASS, we initially planned a 2:1 budgetary ratio for the main sample versus the follow-up sample. The expected cost for the main sample was approximately $10 per unit and, for the follow-up sample, approximately $100 per unit. The follow-up sample response was about 10% of the main sample. As shown in Census comparison Figure 1, inclusion of the non-response sample helped to improve the RASS estimates to more closely mirror the underlying population. While this approach might involve substantially reducing the volume of the main sample for small projects, a balance should still be achieved between the two survey components to maximize total respondents while including some means of assessing non-response bias.

CONCLUSIONS

When reviewing the final RASS results, it was clear that the non-response survey was successful in reaching segments of the population that were underrepresented in the main survey, namely renters, Hispanics, and homes without seniors. When combined with the main sample, the results more closely mirrored the demographic profile of the state based on 2000 Census data. The California RASS experience provides a successful example of how to balance the inherent tradeoffs between volume and accuracy of responses cost-effectively. Researchers undertaking similar surveys should consider the following when designing their sampling approach:

- If flexibility and speed are priorities, a phone survey may be optimal; for cost and high-volume, mail; and for accuracy and detail, on-site
- Select a survey approach based on the time frame of the project and the desired size of the sample
- Specify the sample design to cost-effectively ensure adequate precision for small market segments
- Look at similar prior studies in the target region to determine the expected response rate overall and by population segment
- Attempt to incorporate population-level information into the sample design, e.g., stratification scheme, to address expected variances in response rate across segments
- Consider the tradeoffs associated with implementing a non-response follow-up effort, i.e., volume of the main survey versus the accuracy of the study results
- Compare results with Census or other population-level demographic statistics to determine the extent of non-response bias
- If fielding a follow-up survey, set the sample size large enough to allow for meaningful comparisons with the main survey and the ability to influence the overall study estimates
- Attempt to use multiple survey methods for reaching non-responders (e.g., priority mail with higher incentive, phone and on-site)
- Carefully consider the effect various weighting schemes will have on the overall study precision when weighting the follow-up sample.

FURTHER INFORMATION ON THE PROJECT

The data presented in this paper as well as copies of the reports associated with this study are available on the California RASS web site, which is available to all who register at: http://websafe.kemaine.com/RASSWEB/DesktopDefault.aspx. The project included an interactive web site designed to allow more flexible data review and reporting. The web site includes a process for filtering data as well as producing cross tabulations for individual sections of survey questions or specific survey responses.

REFERENCES

1. Census Data Source: Census 2000 5% PUMS for California
2. Sampling Techniques, William G. Cochran, 1997.