CONSULTANT REPORT

2009 CALIFORNIA RESIDENTIAL APPLIANCE SATURATION STUDY

Volume 1: Methodology

Prepared for: California Energy Commission

Prepared by: KEMA, Inc.

OCTOBER 2010

CEC-200- 2010-004

SB_GT&S_0771146

Prepared by:

Primary Author(s):
KEMA, Incorporated
Claire Palmgren, Project Manager
Noel Stevens
Miriam Goldberg, Ph.D.
Rich Barnes
Karen Rothkin, Ph.D.

KEMA, Inc.
155 Grand Avenue, Suite 500
Oakland, CA 94612
www.kema.com

Contract Number: 400-07-031

Prepared for:

California Energy Commission

Glen Sharp
Gary Occhiuzzo
Contract Managers

Glen Sharp
Project Manager

Bill Junker
Office Manager
Demand Analysis Office

Sylvia Bender
Deputy Director
Electricity Supply Analysis Division

Melissa Jones
Executive Director

ACKNOWLEDGEMENTS

Los Angeles Department of Water and Power (LADWP)

- Michael Cockayne
- Jaimi Johnson

Pacific Gas and Electric (PG&E)

- Rafael Friedmann

San Diego Gas & Electric (SDG&E)

- Eric Drabb

Southern California Edison (SCE)

- Reginald Wilkins

Southern California Gas Company (SoCal Gas)

- George Katsufrakis

ABSTRACT

In 2009, the California Energy Commission funded and administered a Residential Appliance Saturation Study that serves as an update to the 2003 RASS, with the same utilities participating – Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE), San Diego Gas & Electric Company (SDG&E), Southern California Gas Company (SoCal Gas), and Los Angeles Department of Water and Power (LADWP). KEMA was the prime consultant.

The study was implemented as a mail survey with an option for respondents to complete it on-line. The survey requested households to provide information on appliances, equipment, and general consumption patterns. Data collection was completed in early 2010.

The study yielded energy consumption estimates for 27 electric and 10 natural gas residential end-uses and appliance saturations for households. These consumption estimates were developed using a conditional demand analysis, an approach that applied statistical methods to combine survey data, household energy consumption data and weather information to calculate average annual consumption estimates per appliance. The 2009 RASS resulted in end-use saturations for 24,464 individually metered and 1,257 master-metered households. Survey and conditional demand analysis results were weighted to provide population level estimates representative of the participating utilities that allow comparison across utility service territories, forecast climate zones and other variables of interest- dwelling type, dwelling age group, and income.

Keywords: California Energy Commission, conditional demand analysis, CDA, unit energy consumption, UEC, residential, appliance, saturations, degree day normalization, energy survey, data collection

Please use the following citation for this report:

KEMA, Inc. 2010. *2009 California Residential Appliance Saturation Study.* California Energy Commission. Publication number: CEC- 200-2010-004.

TABLE OF CONTENTS

Acknowledgements i

ABSTRACT iii

TABLE OF CONTENTS v

List of Figures viii

List of Tables viii

List of Equations xi

CHAPTER 1: RASS Methodology Introduction 1

CHAPTER 2: Study Design and Implementation 3

Sampling Approach 4

Individually Metered Sample Design 5

Master-Metered Sample Design 13

Non-Response Follow-Up Sample Design 17

Project Implementation 18

Materials Design and Pretest 18

Internet Survey Option 21

Direct Mailings 21

Toll-Free Customer Support 21

Individually Metered Survey Completes 22

Master-Metered Mail Implementation 28

Master-Metered Survey Completes 28

Non-Response Follow-Up Implementation 30

Survey Weights 34

Individually Metered Sample Weights 34

Master-Metered Sample Weights 41

Comparison of Results Across Sampling and Study Groups 42

Non-Response Follow-Up Comparison 42

Online Survey Comparison 43

Master-Metered Comparison 43
Energy Consumption Comparison...44
Precision of RASS Estimates...45
Individually Metered Sample Precision 45
Master-Metered Sample Precision..46
CHAPTER 3: Database Preparation...47
Database 1: Raw RASS Survey Database...48
Database 2: Cleaned RASS Survey and CDA Database 48
Non-Response Indicator..49
Logical Response Inconsistencies 49
Imputing Missing Values..49
Refining Fuel and System Types..49
Estimated UECs..49
Database 3: Billing and Degree-Day Normalization Data File Set.....................................50
Database Formats..50
Data Delivery...50
CHAPTER 4: Data Cleaning And Processing 52
Overview 52
Paper Survey Processing...52
Combined Survey Dataset 52
Excess Responses (G of Figure 4-1)...53
Incomplete Surveys (H of Figure 4-1) 53
Cleaning Individual Survey Questions...56
Invalid Surveys...57
Survey Specific Cleaning..57
Your Home and Lifestyle...58
Additional Cleaning of Your Home and Lifestyle Variables...67
Cleaning Space Heating Survey Responses 68

SB_GT&S_0771153

Space Cooling 71
Cleaning Water Heating Survey Responses 73
Laundry 75
Food Preparation 76
Refrigerators 76
Freezers 77
Spas and Hot Tubs 78
Swimming Pools 78
Entertainment and Technology 79
Lighting 79
Miscellaneous Appliances 80
Household Information 81
Billing and Weather Data 82
Billing Databases 82
Billing Data Cleaning and Preparation for Analysis (Weather-Normalization) 85
Weather-Normalization of Billing Data 87
CDA Variables and Data Imputation Process 95
Advanced Variable Imputation Process 96
Correcting for Fuel Misreporting 98
Space Heating 98
Binary and Continuous CDA Variables 103
Energy Consumption for Refrigerators and Freezers 107
CHAPTER 5: Data Analysis Methodology 109
Statistically Adjusted Engineering Analysis 109
Derivation of End-Use Engineering Estimates 111
Electric End-Use Engineering Estimates 112
Natural Gas End-Use Engineering Estimates 121
Specification of CDA Models 124

Derivation of the Electric SAE 125
Summary of SAE Electric Model 128
Derivation of the Natural Gas SAE 128
Summary of SAE Natural Gas Model 129
Estimated Model Results 129
List of Acronyms 131

LIST OF FIGURES

Figure 2-1: Overview of RASS Study 3
Figure 3-1: Overview of RASS Database Preparation 47
Figure 4-1: Data Cleaning Process Overview 53
Figure 4-2: Comparison of NAC Results and Actual Consumption 95
Figure 5-1: Overview of SAE Process 111

LIST OF TABLES

Table 2-1: Utility Billing Data Requirements 4
Table 2- 2A: Individually Metered Sample Design (PG&E, LADWP and SDG&E) 7
Table 2-2B: Individually Metered Sample Design (SCE) 9
Table 2-3: Master-Metered Sample Design 16
Table 2-4A: Individually Metered Survey Response (PG&E, LADWP and SDG&E) 23
Table 2-4B: Individually Metered Survey Response (SCE) 24
Table 2-5: Master-Metered Phone Survey Completes and Mail-out 29
Table 2-6: Master-Metered Mail Survey Response 30
Table 2-7: Non-Response Completes by Utility and Mode 33
Table 2-8: Response Rates Within ZIP Code Clusters 34
Table 2-9A: Individually Metered Weights (PG&E, LADWP, and SDG&E) 36

Table 2-9B: Individually Metered Weights (SCE) 37
Table 2-10: Master-Metered Weights 41
Table 2-11: Comparison by Surveying Method 42
Table 2-12: Comparison by Response Method 43
Table 2-13: Comparison of Individually Metered to Master-Metered Household Results 44
Table 2-14: Comparison of Energy Consumption for RASS Respondents and Target Population 45
Table 2-15: Precision of Estimates for the Individually Metered Sample 45
Table 2-16: Precision of Estimates for the Master-Metered Sample 46
Table 4-1: Survey Variables Used to Identify Incomplete Surveys 54
Table 4-2: Composite Indicator Variables Used to Identify Incomplete Surveys 55
Table 4-3: Summary of Invalid Surveys 57
Table 4-4: Dwelling Type Cleaning Results for Individually Metered Households 62
Table 4-5: Dwelling Type Cleaning Results for Master Metered Households 63
Table 4-6: Cleaning Results for Missing Home and Lifestyle Responses 64
Table 4-7: Heating Payment Question Cleaning 69
Table 4-8: Primary Heating Fuel Data Cleaning 70
Table 4-9: Cooling Payment Question Cleaning 71
Table 4-10: Water Heating Payment Question Cleaning 73
Table 4-11: Water Heating Fuel Data Cleaning 75
Table 4-12: Missing Refrigerator Number and Characteristics 77
Table 4-13: Missing Freezer Number and Characteristics 78
Table 4-14: Missing Entertainment and Technology Appliances 79
Table 4-15: Missing Miscellaneous Appliances 81
Table 4-16: Missing Household Information 81
Table 4-17: Comparison of Gas and Electric Utility Providers for Survey Respondents 85
Table 4-18: Comparison of Gas and Electric Utility Providers for Survey Respondents With Individually Metered Gas and Electric Service 85
Table 4-19: T24 Climate Zones 88

Table 4-20: Number of Households With Electric DDN Models by T24 Zone and Utility 91
Table 4-21: Number of Households With Gas DDN Models by T24 Zone and Utility..................91
Table 4-22 Best Fit DDN Electric Model Type by T24 Zone..92
Table 4-23: Average Degree-Day Reference Temperature (°F) by Electric Best-Fit Model Type .93
Table 4-24: Best-Fit DDN Gas Model Type by T24 Zone...94
Table 4-25: Primary Space Heating Fuel Cleaning 99
Table 4-26: Auxiliary Space Heating Fuel Cleaning...100
Table 4-27: Primary Water Heating Fuel Cleaning 102
Table 5-1: Space Heating Engineering Estimates..114
Table 5-2: Space Cooling Engineering Estimates..115
Table 5-3: Water Heating Engineering Estimates 116
Table 5-4: Refrigerator and Freezer Engineering Estimates..116
Table 5-5: Kitchen Appliance Engineering Estimates 117
Table 5-6: Laundry Engineering Estimates...118
Table 5-7: Outdoor Lighting Engineering Estimates...118
Table 5-8: Home Electronics and Office Equipment Engineering Estimates...............................119
Table 5-9: Swimming Pool and Spa Engineering Estimates 119
Table 5-10: Well-Pump Engineering Estimates..120
Table 5-11: Forced-Air Fan...120
Table 5-12: Miscellaneous Engineering Estimates..121
Table 5-13: Space Heating Engineering Estimates..122
Table 5-14: Water Heating Engineering Estimates 123
Table 5-15: Range or Oven Engineering Estimates 123
Table 5-16: Clothes Dryer Engineering Estimates 123
Table 5-17: Swimming Pool and Spa Engineering Estimates 124
Table 5-18: Miscellaneous Engineering Estimates..124
Table 5-19: Electric End-Use Terms for SAE Model 127
Table 5-20: Electric SAE Model 130

Table 5-21: Gas SAE Model....130

LIST OF EQUATIONS

Equation 4-1: The DDN Heating and Cooling Model....89

Equation 5-1: General Form of SAE Model....109

Equation 5-2: General Form of Primary Space Heating....113

Equation 5-3: Electric Space Heating....125

Equation 5-4: Electric Space Heating With No Heating Term From the DDN....125

Equation 5-5: Electric Auxiliary Space Heating....125

Equation 5-6: Central Air Conditioning....126

Equation 5-7: Central Air Conditioning With No Cooling Term From the DDN....126

Equation 5-8: Room Air Conditioning....126

Equation 5-9: Evaporative Cooling....126

Equation 5-10: Electric Water Heating....126

Equation 5-11: Electric Kitchen Appliances....126

Equation 5-12: Electric Laundry Equipment....127

Equation 5-13: Spa Filter Pumps and Electric Spa Heating....127

Equation 5-14: Electric Miscellaneous and Ventilation....127

Equation 5-15: SAE Electric Model....128

Equation 5-16: Natural Gas Primary and Auxiliary Space Heating....128

Equation 5-17: Natural Gas Space Heating with Heating Load From DDN....128

Equation 5-18: Natural Gas Water Heating....129

Equation 5-19: Natural Gas Base Load....129

Equation 5-20: SAE Natural Gas Model....129

CHAPTER 1: RASS Methodology Introduction

In 2009, the California Energy Commission funded and administered a Residential Appliance Saturation Study (RASS) that was implemented across the territories of the large investor-owned utilities (IOUs). The 2009 study served as an update to the 2003 RASS, with the same utilities participating – Pacific Gas and Electric Company (PG&E), Southern California Edison Company (SCE), San Diego Gas & Electric Company (SDG&E), Southern California Gas Company (SoCal Gas), and Los Angeles Department of Water and Power (LADWP). KEMA was the prime consultant.

The research team initiated the study in 2008, with the sampling plans and implementation beginning in the spring of 2009. Data was collected using a two-stage direct mail approach to a representative sample of Californian households. The survey requested households to provide information on appliances, equipment, and general usage patterns. The 2003 RASS survey instrument was updated to reflect changes in available energy-consuming technologies in households. An online version of the survey was also developed. A non-response follow-up was implemented after the initial double mailing stage to a sample of the non-respondents. The non-response effort consisted of telephone calls and in-person assistance with completing the survey. Data collection was completed in early 2010.

Survey results were combined with electric and gas billing data provided by each of the participating utilities to model end uses and to calculate estimates of unit energy consumption (UECs) for each electric and natural gas end use. The combined database was used to develop the conditional demand analysis (CDA) using a statistically adjusted engineering model (SAE) approach. The SAE model applied the 2003 RASS CDA formulas to the current survey data, which provided initial engineering estimates for each end use. Normalized annual consumption (NAC) estimates were developed from billing data using a degree-day normalization (DDN) technique. The engineering estimates from each household were regressed against the respective NAC estimates to provide scalar adjustments to the engineering estimates, which were used to estimate new UECs.

The 2009 RASS resulted in enduse saturations for 24,464 individually metered and 1,257 master-metered households. UEC estimates were provided for individually metered households only, while end-use saturations reflected both individually and master-metered households. Survey and CDA results were weighted to provide population level estimates representative of the participating utilities that allow comparison across utility service territories, climate zones, and other variables of interest-dwelling type, dwelling age group, and income, for example.

By using a statewide survey instrument, the research team 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. The Commission-funded sample included sufficient data for utility-specific analyses, but SCE and SDG&E each sponsored the sampling of additional households within their respective service areas that provided them with supplemental data. The project required a joint effort among the study partners, as they collaborated on a research plan, program materials, and implementation strategy. Each utility provided the data necessary to create a unified sampling plan, as well as household-specific information for households that were selected for the sample. The research team provided anonymity to survey participants by assigning a generic identification code that represented the sampling stratification variables. Each participating utility was provided a key to the identification code that allowed the utilities to link survey respondents to a specific account.

Because the study was designed to support interests of a variety of users, the final report included a collection of research products:

- Executive Summary – Presents a summary of key findings.
- Volume One – Describes the study design and implementation methods, along with a detailed description of the data cleaning process and CDA methodology.
- Volume Two – Provides a brief description of the CDA along with tabulated results for end-use UECs and saturations.
- Appendices – All referenced appendices have been compiled into one document for convenience.
- RASS Website – Updated version of the 2003 Internet tool that supports customized queries of the survey data, including the ability to compare 2009 results to 2003 results.

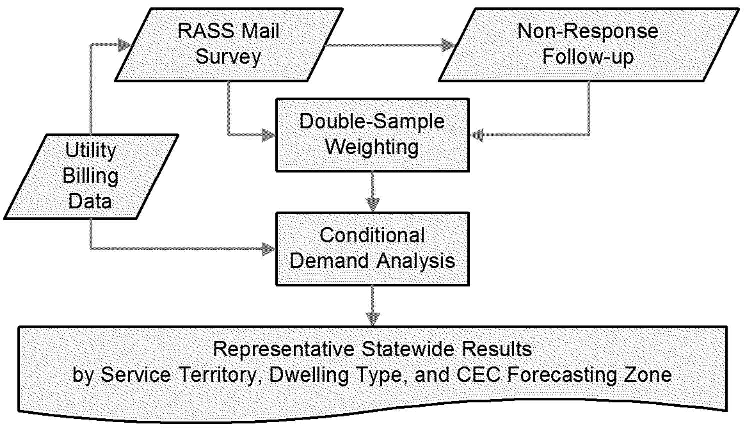
SB_GT&S_0771161

CHAPTER 2: Study Design and Implementation

The objective of the RASS Study was to generate appliance and equipment saturations, general usage patterns, and unit energy consumption (UEC) tables for a set of electric and gas enduses. Survey data was combined with billing data to produce representative results for households in California. An overview of the RASS Study is shown in Figure 2-1.

Participating utilities submitted billing data for their residential population from which a sample of households was selected. Households were invited by mail to complete a RASS survey. To reduce the non-response bias, a subset of households that had not responded to the initial mailing was sent another copy of the survey with an incentive. The survey data were weighted to the population, resulting in a representative database of appliance and equipment holdings and general usage patterns. The survey responses were also matched to the respondent's electric and gas billing data to estimate UEC using conditional demand analysis. The RASS data and conditional demand analysis (CDA) provide saturation and end-use shares that are statistically reliable for service territory, dwelling type, and climate zone.

Figure 2-1: Overview of RASS Study



Source: *2010 California Residential Appliance Saturation Survey*

The RASS study implemented multiple data collection methods to solicit participation from both individually metered and master-metered households. The individually metered households were sent the survey directly. Master-metered accounts that serve between two and four units were surveyed similar to the individually metered but were instructed to fill out the survey for only one unit in the building. Master-metered accounts serving five or more units were surveyed in two stages. The first stage involved a telephone interview with the contact at the complex to acquire information about common-area equipment and to obtain information on mailing addresses for specific households served by the account. The second stage involved mailing a survey to a sample of specific households that were identified in Stage 1.

The Energy Commission sponsored a sample of approximately 100,000 individually metered households plus 5,000 master-metered households. Two participating utilities contributed funding for surveys to be sent to additional individually metered and master-metered households and for their results to be combined with the Energy Commission's RASS results. Southern California Edison (SCE) requested approximately 15,000, and San Diego Gas & Electric (SDG&E) asked for about 12,500 additional surveys to be mailed.

This section discusses the sample design for each of the surveying components, the implementation of the data collection methods, the process of weighting survey results to the population, and the study's resulting precision.

Sampling Approach

The sampling approach for the 2009 RASS followed the general method implemented for the 2003 RASS. The research team based the sample design on the residential population sample frames acquired from the participating electric utilities (PG&E, SCE, LADWP, and SDG&E). Data from participating gas utilities (PG&E, SDG&E, and SoCal Gas) were incorporated into the analysis segment of the study. The utility data required for the study, as outlined in Table 2-1, were transferred in three batches:

Batch 1: Transfer of data required to develop the sample frame

Batch 2: Transfer of data required to mail the surveys

Batch 3: Transfer of data required to process the survey data and perform the conditional demand analysis

Table 2-1: Utility Billing Data Requirements

Electric Population Frame Data

- Premise/Control #
- Service city and 5-digit service ZIP code
- Average Daily kWh consumption for premise over previous 12 months
- Dwelling type indicator
- Geo-demographic indicator (if available)
- Electric rate schedule with baseline allowance codes
- Gas service indicator (PG&E and SDG&E)
- Other geographic indicators (division, forecast climate zone)
- Service description field
- Meter set date or Premise establishment date
- Customer service start date

Contact information for sampled accounts

- Service address
- Mail address, city, state, and ZIP
- Customer name

SoCal Gas Population Frame, and PG&E gas-only accounts or SDG&E gas-only accounts where necessary to match specific accounts in the sample frame

- Premise/control #
- Service address, city and 5-digit service ZIP code
- Mail address, city, state, and ZIP (for supplemental matching)
- Customer name (for supplemental matching)

Transaction billing data (for all sampled accounts)
(One year of transactions for customer that resides in the dwelling at the time of the first survey mailing).

- Kilowatt hour (kWh) and/or therm usage
- Current and prior read dates
- Transaction type (regular bill, adjustment, estimate)
- Applicable rate schedule with baseline allowance
- Any information describing unique characteristics of account

The subsections that follow discuss the sample design for individually metered households, master-metered households, and the non-response follow-up for individually metered households.

Individually Metered Sample Design

The research team implemented a stratified random sample design for individually metered households. The stratification variables for the Energy Commission-sponsored sample and the SDG&E oversample were the same as what had been used for the 2003 RASS: electric utility, age of home, presence of electric heat, home type, and Energy Commission forecast climate zone. The population of individually metered households from PG&E, SCE, SDG&E, and LADWP was distributed across 94 strata. The study team worked with SCE to develop a sampling strategy for its oversample that incorporated T24 zones. The additional SCE households sampled incorporated an oversample of underrepresented combinations of the ENERGY COMMISSION strata with T24 zones. The entire study sample was randomly selected from their respective strata at the same time.

Stratification Variables

The participating electric utilities provided their initial population data, which included four of the five stratification variables. Meter set date was used as a proxy for home age, except for households in PG&E territory that had smart meters installed. The establishment of a new meter set date for the installation of a smart meter precluded its use to indicate home age, so separate strata were constructed for these households. Presence of electric heat was determined by a flag from the utility data. This flag is often used by utilities to determine a household's baseline electricity usage allotment. Home type used a combination of dwelling type and electricity usage for PG&E and SCE households. Dwelling type was not available from SDG&E or LADWP, so only energy usage was used to define home type strata from those two utilities. The Energy Commission forecast climate zone was defined by the service ZIP of the household matched to a climate zone via a lookup table. The additional stratification by T24 zone for the SCE oversample was provided by SCE.

The stratification variables are listed below with their assigned values. Each of the variables became part of the SFCode that was used to identify the stratum per household. The first three positions of the SFCode indicated the electric utility and the stratum from within that utility. The remaining positions of the SFCode, as listed below, represented the specific value of the stratification variable.

- Electric utility (1st position of SFCode): PG&E [P], SCE [S], SDG&E [G for Energy Commission-sponsored or V for oversample], or LADWP [L];
- Age of home (4th position of SFCode): old (prior to 2003) [0] or new (2003 or newer) [1], smart meter (PG&E only) [3];
- Presence of electric heat (5th position in SFCode): yes [1] or no [0];
- Home type (6th position in SFCode):
    - for PG&E and SCE: single family "low" (< 15 kWh/day) [2], single family "high" (> 15 kWh/day) [1], or multifamily (all) [3];
    - for SDG&E and LADWP: high (> 20 kWh/day) [6], medium (10–20 kWh/day) [5], or low (< 10 kWh/day) [4];
- Energy Commission forecast climate zone (7th and 8th positions in SFCode);
- T24 zone (9th and 10th positions in SFCode): SCE only

Sample Frame

The individually metered sample frame is presented in Table 2-2A and Table 2-2B. The columns of the tables are as follows:

- Columns A through E indicate the stratification variables. For Table 2-2B, E2 shows the T24 zone for SCE.
- Column F contains the prefix used for the SFCode in the database to indicate the stratum per respondent.
- Column G shows the proportion of the population for each stratum, where the total population comprises households in PG&E, SCE, SDG&E, and LADWP electric service territories.

- Column H shows the target number of completes for the Energy Commission-sponsored sample using a modified proportional allocation method.
- Column I shows the target number of completes for the SCE and SDG&E oversamples.
- Column J gives the expected response rates per stratum.
- Column K contains the target mail-out, which was determined by the expected response rates along with the target number of completes.
- Column L contains the actual mail-out.

Table 2- 2A: Individually Metered Sample Design (PG&E, LADWP, and SDG&E)

A	B	C	D	E	F	G	H	I	J	K	L
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	Proportion	CEC Target Completes	Over-sample Target Completes	Expected Response Rate	Total Target Mail-out	Total Actual Mail-out
PG&E	OLD	NO	SF HIGH	1	P1	0.64%	122	-	26%	471	465
				2	P2	1.20%	230	-	26%	883	868
				3	P3	2.63%	502	-	26%	1932	1,896
				4	P4	5.63%	1,076	-	26%	4139	4,108
				5	P5	2.90%	554	-	26%	2131	2,110
			SF LOW	1	P6	0.46%	88	-	25%	351	346
				2	P7	0.44%	84	-	25%	336	323
				3	P8	1.01%	193	-	25%	770	744
				4	P9	3.17%	607	-	25%	2426	2,387
				5	P10	2.89%	552	-	25%	2206	2,176
			MF	3	P11	0.84%	161	-	17%	949	912
				4	P12	2.52%	481	-	17%	2827	2,768
				5	P13	3.45%	660	-	17%	3881	3,781
				ALL Other	P14	0.55%	106	-	17%	621	602
		YES	SF HIGH	1	P15	0.81%	155	-	26%	598	590
				2	P16	0.30%	60	-	26%	231	228
				3	P17	0.64%	122	-	26%	469	462
				4	P18	0.49%	93	-	26%	357	352
				5	P19	0.22%	60	-	26%	231	227
			SF LOW	1	P20	0.34%	65	-	29%	224	222
				4	P21	0.22%	60	-	29%	207	203
				5	P22	0.18%	60	-	29%	207	204
				ALL Other	P23	0.22%	60	-	29%	207	201
			MF	4	P24	1.09%	208	-	14%	1485	1,463
				5	P25	1.15%	220	-	14%	1574	1,545
				ALL Other	P26	0.28%	60	-	14%	429	421
	NEW	NO	SF HIGH	2	P27	0.24%	60	-	21%	286	271
				3	P28	0.40%	77	-	21%	367	358
				4	P29	0.40%	76	-	21%	364	355

A	B	C	D	E	F	G	H	I	J	K	L
Electric Utility	**Home Age**	**Electric Heat Presence**	**Home Type**	**CEC Forecast Climate Zone**	**Strata**	**Proportion**	**CEC Target Completes**	**Over-sample Target Completes**	**Expected Response Rate**	**Total Target Mail-out**	**Total Actual Mail-out**
				ALL Other	P30	0.18%	60	-	21%	286	273
			SF LOW	ALL	P31	0.65%	125	-	20%	625	610
			MF	ALL	P32	0.66%	125	-	13%	965	941
		YES	ALL	ALL	P33	0.34%	64	-	22%	292	282
	Smart Meter	NO	SF HIGH	2	P34	0.28%	60	-	22%	273	271
				3	P35	1.11%	212	-	22%	964	938
				ALL Other	P36	0.14%	60	-	22%	273	270
			SF LOW	ALL	P37	0.63%	120	-	22%	547	530
			MF	ALL	P38	0.49%	93	-	22%	422	394
		YES	ALL	ALL	P39	0.14%	60	-	22%	273	264
LADWP	OLD	NO	L	11	L1	3.21%	613	-	14%	4379	4,289
				12	L2	0.96%	183	-	14%	1307	1,267
			M	11	L3	2.04%	390	-	17%	2295	2,262
				12	L4	1.29%	247	-	17%	1452	1,435
			H	11	L5	0.99%	189	-	17%	1109	1,094
				12	L6	1.49%	284	-	17%	1672	1,652
	NEW		L	11	L7	0.55%	105	-	12%	876	851
				12	L8	0.13%	60	-	12%	500	481
			M	11	L9	0.29%	60	-	12%	500	488
				12	L10	0.14%	60	-	12%	500	488
			H	11	L11	0.13%	60	-	12%	500	491
				12	L12	0.15%	60	-	12%	500	490
	ALL	YES	ALL	ALL	L13	0.09%	60	-	22%	273	269
SDG&E	OLD	NO	L	13	G1	2.88%	551	785	23%	5744	5709
			M	13	G2	3.82%	730	1039	26%	6756	6722
			H	13	G3	2.77%	529	753	25%	5127	5112
		YES	L	13	G4	0.06%	60	17	21%	369	368
			M	13	G5	0.16%	60	44	32%	321	318
			H	13	G6	0.29%	60	80	32%	433	432
	NEW	NO	M	13	G7	0.22%	60	61	20%	601	593
			H	13	G8	0.20%	60	54	20%	568	565
		All Other	All Other	13	G9	0.18%	60	49	20%	541	535

Source: *2010 California Residential Appliance Saturation Survey*

Table 2-2B: Individually Metered Sample Design (SCE)

A	B	C	D	E.1	E.2	F	G	H	I	J	K	L
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	T24 Zone	Strata	Proportion	CEC Target Completes	Over-sample Target Completes	Expected Response Rate	Total Target Mail-ou	Total Actual Mail-out
SCE	OLD	NO	SF HIGH	7	13	S1	0.686%	131	80	21%	1,004	995
					14		0.055%	11	2		60	59
					16		0.126%	24	23		226	225
				8	6	S2	1.807%	345	18	21%	1,730	1,727
					8		2.136%	408	14		2,010	1,997
					9,10		0.864%	165	5		811	809
					5		0.028%	5	31		175	174
				9	6	S3	0.227%	43	2	21%	217	217
					8		1.006%	192	7		947	942
					9		2.900%	554	17		2,720	2,711
					14		0.495%	95	20		554	531
					16		0.062%	12	12		112	112
				10	10	S4	3.274%	626	19	21%	3,068	3,030
					14		0.597%	114	24		656	634
					15		0.514%	98	86		876	870
					16		0.229%	44	43		411	407
				Other	6	S5	0.138%	26	1	21%	132	132
					8		0.240%	46	2		226	226
					9		0.002%	-	-		2	2
					10,14, 15,16		0.002%	-	-		2	2
					99		0.001%	-	1		6	5
			SF LOW	7	13	S6	0.249%	47	163	23%	914	898
					14		0.025%	5	5		42	40
					16		0.104%	20	45		282	276
				8	6	S7	1.394%	266	34	23%	1,307	1,298
					8		1.329%	254	22		1,202	1,196
					9,10		0.394%	75	9		367	361
					5		0.031%	6	34		175	174
				9	6	S8	0.257%	49	6	23%	240	237
					8		1.113%	213	19		1,006	999
					9		1.472%	281	14		1,371	1,364
					14		0.184%	35	34		300	292
					16		0.037%	7	16		101	98
				10	10	S9	1.112%	212	36	23%	1,081	1,044

A	B	C	D	E.1	E.2	F	G	H	I	J	K	L
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	T24 Zone	Strata	Proportion	CEC Target Completes	Over-sample Target Completes	Expected Response Rate	Total Target Mail-ou	Total Actual Mail-out
[redacted]					14		0.285%	55	53		467	452
					15		0.160%	31	161		834	816
					16		0.232%	44	101		633	618
				Other	6	S10	0.078%	30	2	23%	138	138
					8,9,10, 14,16		0.078%	29	1		139	133
					99		0.001%	-	1		4	4
			MF	7	13	S11	0.195%	37	167	16%	1,275	1,255
					14		0.020%	4	6		63	59
					16		0.099%	19	72		566	553
				8	6	S12	1.308%	250	22	16%	1,699	1,681
					8		1.376%	263	23		1,785	1,763
					9,10		0.348%	66	9		472	467
					5		0.015%	3	16		119	118
				9	6	S13	0.476%	91	8	16%	619	612
					8		0.680%	130	11		882	868
					9		0.996%	190	26		1,353	1,342
					14		0.141%	27	43		436	420
					16		0.051%	10	37		291	285
				10	10	S14	1.018%	194	34	16%	1,426	1,382
					14		0.177%	34	54		547	523
					15		0.191%	36	158		1,215	1,180
					16		0.066%	13	48		381	365
				Other	6	S15	0.069%	31	1	16%	198	198
					8		0.065%	28	1		185	182
					14		0.001%	-	-		3	3
					9,10, 16		0.001%	-	1		8	7
					99		0.000%	-	-		3	3
		YES	SF HIGH	8	6	S16	0.102%	25	1	20%	132	131
					8		0.084%	21	1		108	108
					9,10		0.050%	13	-		64	64
					5		0.004%	1	4		24	23
				9	6	S17	0.010%	3	-	20%	17	17
					8		0.024%	9	-		44	44
					9		0.072%	26	-		132	132
					14		0.020%	7	1		40	39
					16		0.042%	15	8		114	113
				10	10	S18	0.144%	30	1	20%	154	153
					14		0.074%	15	3		92	90
					15		0.053%	11	9		99	97
					16		0.016%	3	3		32	32

A	B	C	D	E.1	E.2	F	G	H	I	J	K	L
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	T24 Zone	Strata	Proportion	CEC Target Completes	Over-sample Target Completes	Expected Response Rate	Total Target Mail-ou	Total Actual Mail-out
[redacted]				Other	6	S19	0.012%	7	-	20%	35	35
					8		0.007%	4	-		20	20
					9,10		0.000%	-	-		1	1
					13		0.033%	19	4		113	112
					14		0.004%	2	-		13	13
					16		0.027%	15	5		102	102
					99		0.022%	12	19		158	158
			SF LOW	ALL	6	S20	0.078%	15	2	24%	71	70
					8		0.105%	20	2		91	91
					9		0.089%	17	2		80	80
					10		0.042%	8	1		39	36
					13		0.008%	2	5		28	26
					14		0.038%	7	7		59	55
					15		0.031%	6	31		153	150
					16		0.062%	12	27		162	154
					99		0.009%	2	8		42	42
					5		0.002%	-	2		12	12
			MF	8	6	S21	0.682%	130	11	15%	944	939
					8		0.505%	96	8		699	696
					9		0.099%	19	3		143	142
					5		0.009%	2	10		82	82
				9	6	S22	0.156%	30	3	15%	216	215
					8		0.123%	24	2		171	170
					9		0.275%	53	7		398	395
					14		0.030%	6	9		98	94
					16		0.021%	4	15		130	129
				10	10	S23	0.301%	57	10	15%	499	442
					14		0.016%	3	5		52	51
					15		0.063%	12	52		428	414
					16		0.031%	6	23		193	189
				Other	6	S24	0.021%	9	-	15%	60	60
					8		0.013%	6	-		39	39
					13		0.021%	9	18		176	172
					14		0.002%	1	1		12	11
					9,16		0.007%	3	5		50	49
					99		0.078%	33	70		689	682
	NEW	NO	SF HIGH	8	5,6	S25	0.104%	24	2	19%	136	130
					8		0.113%	26	1		142	142
					9,10		0.041%	10	-		51	51
				9	6	S26	0.007%	2	-	19%	9	9
					8		0.016%	4	-		22	22
					9		0.107%	26	1		140	138

A	B	C	D	E.1	E.2	F	G	H	I	J	K	L
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	T24 Zone	Strata	Proportion	CEC Target Completes	Over-sample Target Completes	Expected Response Rate	Total Target Mail-ou	Total Actual Mail-out
[redacted]					14		0.111%	27	4		165	159
					16		0.006%	1	1		14	14
				10	10	S27	0.779%	149	4	19%	806	786
					14		0.211%	40	8		256	254
					15		0.096%	18	16		180	175
					16		0.021%	4	4		41	39
				Other	6,8	S28	0.005%	2	-	19%	10	9
					13		0.120%	44	14		307	302
					14		0.015%	5	1		32	29
					9,16		0.023%	8	4		67	67
			SF LOW	8	6	S29	0.086%	22	2	19%	126	125
					8,9,10		0.149%	38	2		213	184
					5		0.000%	-	1		4	4
				10	10	S30	0.279%	53	9	19%	328	320
					14		0.099%	19	18		195	185
					15		0.029%	5	29		181	172
					16		0.013%	3	6		44	44
				Other	6	S31	0.015%	5	-	19%	26	24
					8		0.022%	7	-		37	37
					9,10		0.051%	15	1		87	86
					13		0.043%	13	28		215	210
					14		0.046%	14	8		117	105
					16		0.022%	7	10		86	82
			MF	Other	5,6	S32	0.068%	13	1	10%	142	138
					8		0.070%	13	1		145	145
					9		0.067%	13	2		147	145
					10		0.083%	16	3		185	181
					13		0.012%	2	10		122	120
					14		0.026%	5	8		130	126
					15		0.013%	2	11		132	130
					16		0.007%	1	5		67	66
		YES	SF HIGH	All	6	S33	0.003%	1	-	17%	8	8
					8		0.003%	1	-		7	7
					9		0.003%	2	-		9	9
					10		0.012%	6	-		33	33
					13		0.003%	1	-		9	9
					14		0.003%	2	-		10	10
					15		0.003%	1	-		11	10
					16		0.002%	1	-		7	7
			SF LOW		6		0.006%	3	-		17	17
					8		0.005%	2	-		16	16
					9		0.005%	2	-		14	14

A	B	C	D	E.1	E.2	F	G	H	I	J	K	L
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	T24 Zone	Strata	Proportion	CEC Target Completes	Over-sample Target Completes	Expected Response Rate	Total Target Mail-ou	Total Actual Mail-out
[redacted]					10		0.012%	5	-		34	33
					13		0.002%	1	1		12	12
					14		0.002%	1	-		8	7
					15		0.002%	1	2		19	19
					16		0.001%	1	1		6	6
					5,6		0.019%	8	-		52	52
					8		0.012%	6	-		33	33
					9		0.003%	2	-		10	10
			MF		10,16		0.023%	10	1		65	65
					13		0.003%	1	3		23	23
					14		0.001%	1	-		5	5
					15		0.003%	1	2		19	18
Total							100.00%	19,942	5,405		125,320	123,332

Source: *2010 California Residential Appliance Saturation Survey*

Initial Mail Sample Allocation

The RASS sample was assigned to each stratum using a modified proportional allocation. A minimum of 60 target completed surveys was set for each stratum. Strata that did not meet the minimum targets, as initially defined by the stratification variables, were combined. Once the minimum sample was assigned to each stratum, the remaining sample points were assigned proportionately to the population distribution.

Even though proportional allocation provides the best precision for the population as a whole or for estimates of saturations or other proportions, assigning target completes per stratum ensures representation in the sample for these strata. Furthermore, the 2003 RASS Study had response rates varying from 8 percent up to 33 percent per stratum. Incorporating the response rates from the previous RASS into the calculation for the minimum mail-out per stratum allowed for larger mail-outs for strata where lower response rates were anticipated.

The target mail-out of the combined Commission-sponsored sample and utility oversamples was set at 125,257, with expected completes set at 25,381, based on an average response rate of 20 percent. The actual mail-out was slightly lower due to accounts having been closed.

Master-Metered Sample Design

The master-metered sample design was constructed by using the electric utility and the type of units the account served. Types of units serviced included master-metered accounts serving two to four units, multifamily complexes with five to twenty units, multifamily complexes with more than twenty units and mobile-home parks with five or more mobile homes. Target completes for the Energy Commission-sponsored sample and the SDG&E oversample were assigned to each type based on the proportion of the population of units within that type (not accounts). The study team worked with SCE to identify segments where they wished to oversample. The study population of master-metered accounts was divided into 16 strata based on electric utility and type of units.

The master-metered accounts were surveyed differently based on what type of units they served. Accounts serving two to four units were surveyed the same way the individually metered households were in that one survey was mailed to the account contact. The cover letter requested the contact to fill out the survey for one unit.

Master-metered accounts serving more than four units were surveyed using a two-step method. The initial step was to conduct a phone interview with the account contact to obtain information about central systems and to request information on specific mailing addresses for the residents. The number of phone calls per stratum was determined by a combination of the number of target-completed surveys and the number of surveys to be mailed out per strata. Complexes with little variation between units (multifamily complexes of four to twenty units) were assigned fewer target complete surveys so that more complexes could be surveyed. Account types with more variation, such as mobile-home parks, were assigned more target complete surveys to capture the variation within complex,

The second step involved selecting households from each complex to send surveys. The number of surveys per complex was determined by the number of units the account served. For multifamily accounts serving five to twenty units, surveys were sent to four households within the complex. For multifamily accounts serving more than twenty units, surveys were sent to eight households within the complex. For mobile-home parks, surveys were sent to ten households within the complex.

Master-Metered Stratification Variables

A stratified random sample design was used for the master-metered accounts. The participating electric utilities provided their initial master-metered population data that included the number of units the account served.

The stratification variables are listed below with their assigned values. The variables became part of the SFCode that was used to identify the stratum per household. The first position of the SFCode indicated the household was master metered and was designated with an "M." The remaining positions of the SFCode, as listed below, represented the specific value of the stratification variable.

- Electric utility (2nd position of SFCode): PG&E [P], SCE [S], SDG&E [G for Commission-sponsored or V for oversample], or LADWP [L]
- Type: 2-4 unit [1], multifamily (5-20 units) [2], multifamily (>20 units) [3], or mobile home (greater than 4 units) [4]

Master-Metered Sample Frame

The master-metered sample frame is presented in Table 2-3. The columns of the table are as follows:

- Columns A and B indicate the strata.
- Column C indicates the SFCode for each stratum.
- Column D shows the total number of units per stratum.
- Column E shows the total number of meters per stratum.
- Column F shows the proportion of the study population of units per stratum.
- Column G shows the target number of phone interviews per stratum.
- Column H shows the target number of mail survey completes per stratum.
- Column I shows the anticipated relative response rate.
- Column J indicates the target mail-outs per stratum that was determined by the expected response rate and the target number of completes.

Table 2-3: Master-Metered Sample Design

A	B	C	D	E	F	G	H	I	J
Utility	Home Type	SFCode Prefix	Number of Units	Number of Meters	Proportion	Target Phone Survey (meters)	Target Survey Completes	Relative Response Rate*	Target Mail-out
PG&E	2-4 units	P1	37,946	17,021	8.83%	n/a	88	20%	440
	Multifamily 5-20 units	P2	17,699	1,971	4.12%	68	41	15%	273
	Multifamily >20 units	P3	43,918	766	10.22%	85	101	15%	680
	Mobile home >4 units	P4	105,112	1,673	24.45%	106	243	23%	1,061
	Subtotal		**204,675**	**21,431**	**47.61%**	**259**	**473**	**20%**	**2,454**
SDG&E	2-4 units	G1	7,694	3,465	1.79%	n/a	39	20%	195
	Multifamily 5-20 units	G2	5,558	542	1.29%	47	29	15%	187
	Multifamily >20 units	G3	9,235	175	2.15%	39	47	15%	307
	Mobile home >4 units	G4	33,469	413	7.78%	73	173	23%	730
	Subtotal		**55,956**	**4,595**	**13.02%**	**159**	**287**	**20%**	**1,419**
SCE	2-4 units	S1	10,789	4,615	2.51%	n/a	52	20%	260
	Multifamily 5-20 units	S2	12,695	1,363	2.95%	101	60	15%	406
	Multifamily >20 units	S3	19,974	350	4.65%	79	95	15%	635
	Mobile home >4 units	S4	107,862	1,500	25.09%	238	543	23%	2,369
	Subtotal		**151,320**	**7,828**	**35.20%**	**418**	**750**	**20%**	**3,670**
LADWP	2-4 units	L1	1,480	615	0.34%	n/a	3	20%	15
	Multifamily 5-20 units	L2	3,042	246	0.71%	12	7	15%	47
	Multifamily >20 units	L3	9,676	176	2.25%	19	22	15%	153
	Mobile home >4 units	L4	3,775	27	0.88%	4	9	23%	39
	Subtotal		**17,973**	**1,064**	**4.18%**	**35**	**42**	**20%**	**254**
TOTAL	2-4 units		57,909	25,716	13.47%	n/a	182	20%	910
	Multifamily 5-20 units		38,994	4,122	9.07%	228	137	15%	913
	Multifamily >20 units		82,803	1,467	19.26%	222	265	15%	1,775
	Mobile home >4 units		250,218	3,613	58.20%	421	968	23%	4,199
	Total		**429,924**	**34,918**	**100.00%**	**871**	**1,552**	**20%**	**7,797**

Source: *2010 California Residential Appliance Saturation Survey*

Master-Metered Initial Mail Sample Allocation

The combined Energy Commission-sponsored and the utility-sponsored RASS sample was assigned proportionately to each stratum, based on units or dwellings.

The target phone surveys for the complete sample was set at 871, and the total number of surveys mailed out was set at 7,797. The response rates from the master-metered households from the 2003 RASS Study were included in the calculation for the target mail-outs. The total target number of completes was 1,552, based on an expected overall average response rate of 20 percent.

Non-Response Follow-Up Sample Design

The objective of the non-response effort of the study was to help reduce non-response bias by obtaining responses from a portion of households that had not responded to the first two mail solicitations. A subset of 6,000 individually metered households was selected from the original RASS sample who had not responded.

The population of non-responders was distributed across 1,436 ZIP codes. The more densely populated areas of the state were clustered for sampling by ZIP code to allow more efficient in-person data collection. Since the more sparsely populated areas of the state were not contacted in-person, the households were sampled by strata instead of ZIP codes.

Sample Selection

- Step 1: Separate households by sample group

The first step in the sample selection was to separate the households into clustered and non-clustered sample groups. The following 3-digit ZIP codes were assigned to the non-clustered sample group:

  - 934, 935, 939, 949, 954, 955, 960 and 961

In addition to the ZIP codes specified above, the selection procedure for the clustered sample resulted in inadequate sample sizes for certain strata. If a stratum had fewer than 246 cases, the stratum was re-assigned to the non-clustered sample group to ensure that at least five cases would be selected from that stratum. Households in the following 25 strata were reassigned to the non-clustered sample group (regardless of their ZIP code):

  - G04, G05, G06
  - L13
  - P01, P06, P16, P17, P18, P19, P20, P21, P22, P23, P27, P29, P30, P33, P34, P36, P39
  - S10, S16, S17, S26

The households not assigned to the non-clustered sample group were assigned to the clustered sample group.

- Step 2: Select the sample members from the non-clustered group

The second step in the sample selection was to select households from within the non-clustered sample group. The non-clustered sample group contained 11,063 records, constituting 12.7 percent of the total non-response households. The non-clustered group was allocated 760 of the 6,000 sample cases and was selected by stratum.

A stratified random sample of households was selected by the following procedure:

  - Allocated sample fraction to strata: A sample fraction was calculated as the total number of cases in a stratum divided by the total number of cases in the non-clustered stratum.

- Allocated sample cases to strata: The sample fraction was multiplied by the sample allocation (760) to get the number of sample cases allocated to each stratum. A statistical rounding procedure provided an integer value for each stratum.
- Cases were randomly selected from each stratum.

The sampling rate for cases selected from the non-clustered sample was 0.0687. The number of cases selected from the 60 strata ranged from 5 to 47 cases.

- Step 3: Select the sample members from the clustered group

The third step in the sample selection was to select households from within the clustered sample group. The clustered ZIP codes contained 75,973 records, comprising 87.3 percent of the total non-response households. The clustered group was allocated 5,240 of the 6,000 sample cases. Households in the clustered group were distributed across 1,147 zip codes and were selected from within ZIP code groups.

The clustered sample was selected by the following procedure:

- Created ZIP code groups: ZIP codes were sorted and grouped to meet a minimum of 20 cases per group. The ZIP codes were collapsed into 816 ZIP code groups.
- Randomly selected 262 zip code groups: The number of ZIP code groups selected was calculated as the number of cases in the sample divided by 20 cases per group.
- = 5240/20=262 ZIP code groups selected
- Randomly selected 20 cases from each of the selected zip code groups. The sampling rate within each ZIP code group was 20 divided by the number of cases within the ZIP code group.

The sampling rate for cases selected from the clustered sample was 0.0690.

Project Implementation

While the agreements with the utilities for data transfer were being negotiated and the sample frame was being developed and finalized, the study team collaborated on updates of the survey materials and planned the overall project implementation. This section details the results of that planning and implementation effort.

Materials Design and Pretest

The materials for the 2009 RASS were based on the materials from the 2003 RASS. All materials were reviewed by the Energy Commission and the participating utilities for content and appearance. All parties agreed on revisions to the cover letters, outer envelope, and survey instrument.

The survey instrument was revised to reflect additional areas of interest and updated technologies. Specific changes to the survey included the following, listed by survey section:

Home and Lifestyle:

- Aggregated year residence built
- Addition of vehicle questions
  - Number of vehicles
  - One-way miles driven on regular weekday trips
  - Total miles per year
  - Parking location
  - Electrical outlet available for parking space

Space Heating:

- Addition of gas fireplace for primary and additional heating
- Gas heating pilot lights split into categories of on all year and on in winter only

Space Cooling:

- Addition of indication of main system controlled with a zoned thermostat
- Addition of programmable communicating thermostat

Water Heating:

- Addition of natural gas high-efficiency condensing water heater
- Deletion of water heater tank insulation

Food Preparation:

- Addition of age of microwave oven
- Changed food preparation usage to list equipment

Refrigerators:

- Moved discard to Miscellaneous Appliances section

Freezers:

- Moved discard to Miscellaneous Appliances section

Spas and Hot Tubs:

- Added time of day per season to running the filter pump
- Added time of day per season to frequency of heating

Swimming Pools:

- Added time of day per season to operating the filter
- Added time of day per season to frequency of heating
- Added pool vacuum as option for pool attributes

Entertainment and Technology:

- Expanded TVs to indicate type
- Updated list of entertainment technologies to reflect current types

- Incorporated size into TV usage
- Split personal computers and usage into desktop and laptop categories
- Added playing games to activities performed on computer
- Updated list of home office equipment to reflect current types

Lighting:

- Added question to indicate number of compact fluorescent (CFL) and incandescent light bulbs per home area
- Added question to indicate proportion of lights used by time of day
- Added nightlights to list of indoor lighting products
- Added question to indicate if replaced a CFL with an incandescent

Miscellaneous Appliances:

- Changes to list of appliances:
  - Added chargers left plugged in all of the time, domestic hot water recirculation pump, and wine or beverage cooler to list of appliances
  - Combined humidifier and dehumidifier into one category of appliance
  - Deleted waterbed from list of appliances
- Added evaporative cooler, refrigerator, and stand-alone freezer to lists of equipment added and equipment discarded

The 2009 survey instrument is contained in Appendices Volume, Appendix A. Lists of variables from the 2009 survey and the 2003 survey are available in Appendices Volume, Appendix C.

The updated materials were pretested with 30 energy customers in San Diego, Anaheim, and Oakland. Results of the pretest were shared with the study team, and final modifications were made to the materials. The results from the materials pretest are in Appendices Volume, Appendix D.

The direct mail solicitation package included the following items:

- An outgoing envelope (7.5 inches by 10.5 inches) with a window opening
- A cover letter: one of the following:
  - Standard first-mailing letter
  - First-mailing letter for sites with 2-4 units
  - First-mailing letter for master-metered sites
  - Second-mailing letter (used for all sites)
- A 20-page scannable survey in English (6.75 inches by 9.75 inches)
- A business-reply envelope (7 inches by 10 inches)

The direct mail materials are included in Appendices Volume, Appendix E.

The survey was also translated into Spanish. Although it was not printed as a scannable form, it was created as a Microsoft Word® file in the same format as the English scannable version. The Spanish version is in the Appendices Volume, Appendix B.

Internet Survey Option

Once the materials were finalized, online versions of the survey instrument were developed in English and Spanish. The online versions of the survey followed the format of the questions from the print version as closely as possible. Customers were directed to the on-line surveys in the cover letter and in the instructions on the inside cover of the paper surveys. Access to the online version required participants to enter their identification code as printed on the survey cover along with the ZIP code of the service address. Participants were given the choice of accessing the English or Spanish version of the survey on the introductory screen. Screen shots of the English online version are in Appendices Volume, Appendix F, and screen shots of the Spanish version are in Appendices Volume, Appendix G.

Direct Mailings

Customer names and mailing addresses were printed on the cover of the surveys in an area where they would show through the window of the outgoing envelope. A bar code, containing a tracking number, and the service address were also printed on the survey cover. Instructions on the inside cover directed respondents to complete the survey for the service address printed on the cover. A cover letter identified study sponsors, provided background information on the study, and gave limited instructions for completing the survey. The survey included instructions on filling out the scannable form.

The direct mail packages were assembled, presorted, and mailed third-class from a mailing house. Direct mail solicitations for the individually metered sample were sent out in batches by utility, beginning on May 18, 2009, through June 5, 2009. The oversample for SCE was mailed on June 18, 2009, and the oversample for SDG&E was mailed later in the year, on November 12, 2009. The bar codes on the surveys were scanned as they arrived and a list was created of the surveys received. Names of respondents who had replied were removed from the list to receive a second solicitation package. The second mailings were also sent out in batches by utility, beginning on June 25, 2009, through July 24, 2009. The second mailing for SDG&E was made on January 13, 2010, to avoid the holiday mail. The cover letter for the second mailing stressed the importance of receiving information from everyone and the outgoing envelopes had an alternate phrase, but all other materials were identical to the first mailing. Data collection protocols are contained in Appendices Volume, Appendix H, including the survey processing steps, training information, and phone scripts for the master-metered electric accounts.

Toll-Free Customer Support

A toll-free telephone support line was provided to assist customers who had questions about the study, how to complete the survey, or issues with the online survey. The phone line also provided support to non-English speaking respondents. Operators had the Spanish version of the survey available to administer over the phone. Several operators were native Spanish speakers.

An operator answered the phone between 8:30 a.m. and 5:00 p.m. Calls received outside those hours were forwarded to voice mail, with an operator retuning calls within a business day or as instructed by the caller.

A total of 562 phone calls were received on the toll-free phone line during the RASS study. There were 361 English calls, 197 Spanish calls, 3 Chinese calls, and 1 Russian call. A translation firm assisted with the Chinese and Russian calls.

Individually Metered Survey Completes

A total of 22,141 completed surveys were received from the first two mailings for the individually metered sample. The counts of completed surveys include paper surveys returned by mail and surveys completed online. The expected response rate was estimated at 20 percent, and the actual response rate for the initial mailing was 18 percent. Table 2-4 presents the number of completed surveys and the response rate for each of the individually metered stratum. The columns of the table include the following information:

- Columns A through E indicate the stratification variables.
- Column F contains the prefix used for the SFCode in the database to indicate the stratum per respondent.
- Column G shows the target number of completed surveys.
- Column H shows the actual number of completes.
- Column I shows the expected response rate.
- Column J gives the actual response rates per stratum.

Table 2-4A: Individually Metered Survey Response (PG&E, LADWP, and SDG&E)

A	B	C	D	E	F	G	H	I	J
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	Target Completes	Actual Completes	Expected Response Rate	Actual Response Rate
PG&E	OLD	NO	SF HIGH	1	P1	122	138	26%	30%
PG&E	OLD	NO	SF HIGH	2	P2	230	189	26%	22%
PG&E	OLD	NO	SF HIGH	3	P3	502	397	26%	21%
PG&E	OLD	NO	SF HIGH	4	P4	1076	935	26%	23%
PG&E	OLD	NO	SF HIGH	5	P5	554	439	26%	21%
PG&E	OLD	NO	SF LOW	1	P6	88	104	25%	30%
PG&E	OLD	NO	SF LOW	2	P7	84	76	25%	24%
PG&E	OLD	NO	SF LOW	3	P8	193	174	25%	23%
PG&E	OLD	NO	SF LOW	4	P9	607	572	25%	24%
PG&E	OLD	NO	SF LOW	5	P10	552	526	25%	24%
PG&E	OLD	NO	MF	3	P11	161	97	17%	11%
PG&E	OLD	NO	MF	4	P12	481	449	17%	16%
PG&E	OLD	NO	MF	5	P13	660	547	17%	14%
PG&E	OLD	NO	MF	ALL Other	P14	106	96	17%	16%
PG&E	OLD	YES	SF HIGH	1	P15	155	176	26%	30%
PG&E	OLD	YES	SF HIGH	2	P16	60	60	26%	26%
PG&E	OLD	YES	SF HIGH	3	P17	122	125	26%	27%
PG&E	OLD	YES	SF HIGH	4	P18	93	95	26%	27%
PG&E	OLD	YES	SF HIGH	5	P19	60	61	26%	27%
PG&E	OLD	YES	SF LOW	1	P20	65	78	29%	35%
PG&E	OLD	YES	SF LOW	4	P21	60	39	29%	19%
PG&E	OLD	YES	SF LOW	5	P22	60	45	29%	22%
PG&E	OLD	YES	SF LOW	ALL Other	P23	60	54	29%	27%
PG&E	OLD	YES	MF	4	P24	208	170	14%	12%
PG&E	OLD	YES	MF	5	P25	220	203	14%	13%
PG&E	OLD	YES	MF	ALL Other	P26	60	57	14%	14%
PG&E	NEW	NO	SF HIGH	2	P27	60	36	21%	13%
PG&E	NEW	NO	SF HIGH	3	P28	77	48	21%	13%
PG&E	NEW	NO	SF HIGH	4	P29	76	47	21%	13%
PG&E	NEW	NO	SF HIGH	ALL Other	P30	60	55	21%	20%
PG&E	NEW	NO	SF LOW	ALL	P31	125	106	20%	17%
PG&E	NEW	NO	MF	ALL	P32	125	91	13%	10%
PG&E	NEW	YES	ALL	ALL	P33	64	40	22%	14%
PG&E	Smart Meter	NO	SF HIGH	2	P34	60	56	22%	21%
PG&E	Smart Meter	NO	SF HIGH	3	P35	212	161	22%	17%
PG&E	Smart Meter	NO	SF HIGH	ALL Other	P36	60	48	22%	18%
PG&E	Smart Meter	NO	SF LOW	ALL	P37	120	98	22%	18%
PG&E	Smart Meter	NO	MF	ALL	P38	93	41	22%	10%
PG&E	Smart Meter	YES	ALL	ALL	P39	60	28	22%	11%
LADWP	OLD	NO	L	11	L1	613	549	14%	13%
LADWP	OLD	NO	L	12	L2	183	152	14%	12%
LADWP	OLD	NO	M	11	L3	390	388	17%	17%
LADWP	OLD	NO	M	12	L4	247	234	17%	16%
LADWP	OLD	NO	H	11	L5	189	213	17%	19%
LADWP	OLD	NO	H	12	L6	284	327	17%	20%

A	B	C	D	E	F	G	H	I	J
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	Target Completes	Actual Completes	Expected Response Rate	Actual Response Rate
	NEW		L	11	L7	105	106	12%	12%
				12	L8	60	34	12%	7%
			M	11	L9	60	48	12%	10%
				12	L10	60	64	12%	13%
			H	11	L11	60	46	12%	9%
				12	L12	60	68	12%	14%
	ALL	YES	ALL	ALL	L13	60	39	22%	14%
SDG&E	OLD	NO	L	13	G1	1336	945	23%	17%
			M	13	G2	1769	1387	26%	21%
			H	13	G3	1282	970	25%	19%
		YES	L	13	G4	77	67	21%	18%
			M	13	G5	104	64	32%	20%
			H	13	G6	140	90	32%	21%
	NEW	NO	M	13	G7	121	74	20%	12%
			H	13	G8	114	75	20%	13%
		All Other	All Other	13	G9	109	67	20%	13%

Source: *2010 California Residential Appliance Saturation Survey*

Table 2-4B: Individually Metered Survey Response (SCE)

A	B	C	D	E	F.1	F.2	G	H	I	J
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Target Completes	Actual Completes	Expected Response Rate	Actual Response Rate
SCE	OLD	NO	SF HIGH	7	S1	13	211	195	21%	20%
						14	13	11		19%
						16	47	63		28%
				8	S2	6	363	374	21%	22%
						8	422	393		20%
						9,10	170	140		17%
						5	36	42		24%
				9	S3	6	45	34	21%	16%
						8	199	135		14%
						9	571	516		19%
						14	115	99		19%
						16	24	23		21%
				10	S4	10	645	593	21%	20%
						14	138	134		21%
						15	184	205		24%
						16	87	78		19%
				Other	S5	6	27	29	21%	22%
						8	48	48		21%
						9	0	1		50%
						10,14,15,16	0	2		100%
						99	1	2		40%
			SF LOW	7	S6	13	210	144	23%	16%
						14	10	5		13%
						16	65	66		24%

A	B	C	D	E	F.1	F.2	G	H	I	J
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Target Completes	Actual Completes	Expected Response Rate	Actual Response Rate
[redacted]				8	S7	6	300	301	235	23%
						8	276	280		23%
						9,10	84	77		21%
						5	40	55		32%
				9	S8	6	55	50	23%	21%
						8	232	153		15%
						9	295	283		21%
						14	69	39		13%
						16	23	24		24%
				10	S9	10	248	198	23%	19%
						14	108	74		16%
						15	192	197		24%
						16	145	143		23%
				Other	S10	6	32	26	23%	19%
						8,9,10,14,16	30	37		28%
						99	1	2		50%
			MF	7	S11	13	204	207	16%	16%
						14	10	8		14%
						16	91	130		24%
				8	S12	6	272	261	16%	16%
						8	286	205		12%
						9,10	75	65		14%
						5	19	18		15%
				9	S13	6	99	63	16%	10%
						8	141	98		11%
						9	216	147		11%
						14	70	56		13%
						16	47	45		16%
				10	S14	10	228	232	16%	17%
						14	88	81		15%
						15	194	284		24%
						16	61	37		10%
				Other	S15	6	32	24	16%	12%
						8	29	22		12%
						14	0	1		33%
						9,10,16	1	2		29%
						99	0	1		33%
		YES	SF HIGH	8	S16	6	26	39	20%	30%
						8	22	25		23%
						9,10	13	9		14%
						5	5	12		52%
				9	S17	6	3	3	20%	18%
						8	9	12		27%
						9	26	24		18%
						14	8	16		41%
						16	23	26		23%
				10	S18	10	31	41	20%	27%
						14	18	24		27%
						15	20	16		16%

A	B	C	D	E	F.1	F.2	G	H	I	J
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Target Completes	Actual Completes	Expected Response Rate	Actual Response Rate
[redacted]						16	6	11		34%
				Other	S19	6	7	7	20%	20%
						8	4	3		15%
						9,10	0	1		100%
						13	23	20		18%
						14	2	5		38%
						16	20	27		26%
						99	31	72		46%
			SF LOW	ALL	S20	6	17	15	24%	21%
						8	22	16		18%
						9	19	11		14%
						10	9	3		8%
						13	7	3		12%
						14	14	17		31%
						15	37	33		22%
						16	39	30		19%
						99	10	14		33%
						5	2	3		25%
			MF	8	S21	6	141	139	15%	15%
						8	104	76		11%
						9	22	20		14%
						5	12	11		13%
				9	S22	6	33	40	15%	19%
						8	26	21		12%
						9	60	45		11%
						14	15	9		10%
						16	19	26		20%
				10	S23	10	67	42	15%	10%
						14	8	12		24%
						15	64	51		12%
						16	29	22		12%
				Other	S24	6	9	5	15%	8%
						8	6	5		13%
						13	27	23		13%
						14	2	4		36%
						9,16	8	13		27%
						99	103	257		38%
	NEW	NO	SF HIGH	8	S25	5,6	26	21	19%	16%
						8	27	13		9%
						9,10	10	10		20%
				9	S26	6	2	2	19%	22%
						8	4	1		5%
						9	27	15		11%
						14	31	23		14%
						16	2	4		29%
				10	S27	10	153	121	19%	15%
						14	48	42		17%
						15	34	26		15%
						16	8	7		18%

A	B	C	D	E	F.1	F.2	G	H	I	J
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Target Completes	Actual Completes	Expected Response Rate	Actual Response Rate
[redacted]				Other	S28	6,8	2	1	19%	11%
						13	58	38		13%
						14	6	6		21%
						9,16	12	20		30%
			SF LOW	8	S29	6	24	11	19%	9%
						8,9,10	40	21		11%
						5	1	1		25%
				10	S30	10	62	40	19%	13%
						14	37	28		15%
						15	34	29		17%
						16	9	6		14%
				Other	S31	6	5	3	19%	13%
						8	7	2		5%
						9,10	16	16		19%
						13	41	26		12%
						14	22	12		11%
						16	17	12		15%
			MF	Other	S32	5,6	14	16	10%	12%
						8	14	19		13%
						9	15	24		17%
						10	19	12		7%
						13	12	20		17%
						14	13	19		15%
						15	13	17		13%
						16	6	11		17%
		YES	SF HIGH	All	S33	6	1	1	17%	13%
						8	1	1		14%
						9	2	0		0%
						10	6	7		21%
						13	1	1		11%
						14	2	0		0%
						15	1	2		20%
						16	1	3		43%
			SF LOW			6	3	3		18%
						8	2	1		6%
						9	2	3		21%
						10	5	5		15%
						13	2	4		33%
						14	1	1		14%
						15	3	3		16%
						16	2	2		33%
			MF			5,6	8	8		15%
						8	6	4		12%
						9	2	3		30%
						10,16	11	6		9%
						13	4	4		17%
						14	1	0		0%
						15	3	3		17%
Total							25,347	22,141	20%	18%

Source: *2010 California Residential Appliance Saturation Survey*

Master-Metered Mail Implementation

The master-metered implementation consisted of two stages. The first stage entailed a phone interview with the contact listed on the account, which was the manager of the facility in many cases. The interviews resulted in obtaining information on common-area equipment and eliciting specific address information for mailing the packet to the households. Phone scripts are contained in Appendices Volume, Appendix H.

Tracking databases of contact information were created for each utility's master-metered sample. Each phone surveyor entered the resident address information from the facility contact into an address database in preparation for creating mailing address files.

Master-metered homes that are in buildings of 2-4 units were not included in the phone survey phase. They received a mail survey packet with a cover letter requesting that they complete the survey for only one unit. Mailing addresses for the other three categories of master-metered home types were based on the information obtained during the phone interviews.

The mailing address files were submitted to the mailing house and printed on the surveys. The information on central systems or community access (swimming pools, spas, etc.) was transferred on to the surveys for the specific households, and the mailing packets were assembled. Prefilling the sections for common equipment with information provided by the facility managers improved the accuracy of responses. The surveys were mailed with a cover letter with instructions to skip questions that had already been filled out.

Master-Metered Survey Completes

A total of 784 phone interviews were completed with 7,397 surveys subsequently mailed to master-metered households. Table 2-5 presents the number of phone interviews conducted and the mail surveys sent by stratum. The columns of the table include:

- Columns A and B indicate the strata.
- Column C indicates the SFCode for each stratum.
- Column D shows the target number of phone surveys and actual completes per stratum.
- Column E shows the target mail-out and actual mail-out per stratum.

For several strata, the target number of phone calls was not reached for a variety of reasons. In some cases, the stratum contained information on a limited number of facilities, and the attempts were exhausted before the targets were met. In other cases, the facility managers were not willing to answer the questions about their facility. The number of phone calls required before completing an interview varied by home type. Mobile-home parks had the lowest number of calls required per complete at 4.5 calls. Multifamily complexes with more than 20 units required 11.9 calls per complete, and multifamily complexes with 5 to 20 units were the most difficult, requiring 16.0 calls per complete.

Table 2-5: Master-Metered Phone Survey Completes and Mail-out

A	B	C	D		E	
			Phone Surveys		Mailout	
Utility	Home Type	SFCode Prefix	Target Phone Survey (meters)	Actual Completes	Target Mailout	Actual Mailout
PGE	2-4 units	MP1	n/a	n/a	440	432
	Multi-family 5-20 units	MP2	68	68	273	272
	Multi-family >20 units	MP3	85	85	680	663
	Mobile home >4 units	MP4	106	106	1,061	1,053
	Sub-Total		**259**	**259**	**2,454**	**2,420**
SDGE	2-4 units	MG1	n/a	n/a	195	192
	Multi-family 5-20 units	MG2	47	22	187	100
	Multi-family >20 units	MG3	39	39	307	288
	Mobile home >4 units	MG4	73	60	730	698
	Sub-Total		**159**	**121**	**1,419**	**1,278**
SCE	2-4 units	MS1	n/a	n/a	260	313
	Multi-family 5-20 units	MS2	101	96	406	368
	Multi-family >20 units	MS3	79	35	635	272
	Mobile home >4 units	MS4	238	238	2,369	2,500
	Sub-Total		**418**	**369**	**3,670**	**3,453**
LADWP	2-4 units	ML1	n/a	n/a	15	14
	Multi-family 5-20 units	ML2	12	12	47	48
	Multi-family >20 units	ML3	19	19	153	144
	Mobile home >4 units	ML4	4	4	39	40
	Sub-Total		**35**	**35**	**254**	**246**
TOTAL	2-4 units		n/a	n/a	910	951
	Multi-family 5-20 units		228	198	913	788
	Multi-family >20 units		222	178	1,775	1,367
	Mobile home >4 units		421	408	4,199	4,291
	Total		**871**	**784**	**7,797**	**7,397**

Source: *2010 California Residential Appliance Saturation Survey*

A total of 1,257 surveys were completed for master-metered households. The expected response rate was 20 percent, and the overall actual response rate was 17 percent. The response rate from households in mobile-home parks was the highest, with the other categories of multifamily households achieving response rates similar to the multifamily strata of the individually metered sample.

Table 2-6 presents the number of mail surveys completed and the response rates for the master-metered households. The columns contain the following information:

- Columns A and B indicate the strata.
- Column C indicates the SFCode for each stratum.
- Column D shows the target number of completed surveys per stratum.
- Column E shows the actual surveys completed per stratum.

- Column F contains the expected response rate per stratum.
- Column G provides the actual response rate per stratum.

Table 2-6: Master-Metered Mail Survey Response

A	B	C	D	E	F	G
			Mail Surveys			
Utility	Home Type	SFCode Prefix	Target Survey Completes	Actual Survey Completes	Expected Response Rate	Actual Response Rate
PGE	2-4 units	MP1	88	72	20%	17%
	Multi-family 5-20 units	MP2	41	35	15%	13%
	Multi-family >20 units	MP3	101	81	15%	12%
	Mobile home >4 units	MP4	243	229	23%	22%
	Sub-Total		**473**	**417**	**20%**	**17%**
SDGE	2-4 units	MG1	39	26	20%	14%
	Multi-family 5-20 units	MG2	29	13	15%	13%
	Multi-family >20 units	MG3	47	12	15%	4%
	Mobile home >4 units	MG4	173	144	23%	21%
	Sub-Total		**287**	**195**	**20%**	**15%**
SCE	2-4 units	MS1	52	49	20%	16%
	Multi-family 5-20 units	MS2	60	15	15%	4%
	Multi-family >20 units	MS3	95	28	15%	10%
	Mobile home >4 units	MS4	543	530	23%	21%
	Sub-Total		**750**	**622**	**20%**	**18%**
LADWP	2-4 units	ML1	3	3	20%	21%
	Multi-family 5-20 units	ML2	7	3	15%	6%
	Multi-family >20 units	ML3	22	7	15%	5%
	Mobile home >4 units	ML4	9	10	23%	25%
	Sub-Total		**42**	**23**	**20%**	**9%**
TOTAL	2-4 units		182	150	20%	16%
	Multi-family 5-20 units		137	66	15%	8%
	Multi-family >20 units		265	128	15%	9%
	Mobile home >4 units		968	913	23%	21%
	Total		**1,552**	**1,257**	**20%**	**17%**

Source: *2010 California Residential Appliance Saturation Survey*

Non-Response Follow-Up Implementation

As described in the sampling section of this chapter, a sample of 5,988 households that had not responded to earlier mailings was selected for the non-response follow-up. The mailing information was provided to the mailing house, and after filtering out problem addresses, 5,671 surveys were mailed to households that had not responded to the survey. The non-response sample was stratified into rural and urban areas and by ZIP code, with 732 (13 percent) rural and 4,939 (87 percent) urban households. A third copy of the 2009 RASS Survey was mailed to the non-response sample, followed by a field effort targeted toward the non-responders in urban areas. The field effort was organized to achieve a minimum of a 50 percent response rate in each of the ZIP code clusters that had been selected in the sample. Field visits were scheduled according to a prioritization of ZIP code clusters with the greatest need for responses. A database was developed to aid with tracking response rates from the field effort. This section details the implementation and results of the non-response follow-up to the initial mailings of the RASS survey.

Non-Response Follow-Up Materials

The priority mail effort was designed based upon experience from the 2003 RASS. For the 2003 RASS, urban households received a first-class mailing with $1 bill and rural households received a priority mailing with $5 bill and a promise of a $15 incentive. Based upon higher response rates achieved with the priority mailer and the higher incentive, priority mail was used for all households in the non-response sample for the 2009 RASS. The priority mailer also contained a $5 bill and a promise of a $10 gift card for returning the survey. The cover letter was targeted toward non-responders.

The priority mailer packets were sent on October 2, 2009. The following materials were included in the non-response mailing package, with examples provided in Appendices Volume, Appendix I.

- Priority mail envelope
- Cover letter
- Survey instrument (same as main sample survey, in Appendix A)
- Card with $5 bill attached and promise of $10 gift card

The field follow-up effort was targeted only at urban households. Postcards were sent in advance of the field visits to alert households in the urban areas that they might be contacted either by phone or in person. Field staff carried badges and letters of introduction from each of the utilities as credentials.

The pre-visit postcards were sent on October 23, 2009. Field visits commenced on October 30, 2009, and were completed by December 12, 2009. The following materials were created for the non-response field effort, with examples provided in Appendices Volume, Appendix J.

- Pre-visit postcard
- Field researcher badge
- Utility specific letters of introduction in English and Spanish
- Postcard to leave behind
- Letter of appreciation to accompany the gift card
- $10 gift cards

Prior to being sent into the field, the field staff was given an orientation training to the RASS survey instrument and was trained in techniques for interacting with customers and responding to the most frequently asked questions.

The following protocols were created for the non-response field effort, provided in Appendices Volume, Appendix K.

- Telephone script
- Field visit script

Non-Response Follow-Up Data Collection

Within a week of the priority mailer going out, completed surveys from the non-response sample began arriving. The protocol for handling the surveys was similar to that used for the initial mailings. The surveys were logged and the bar codes scanned. The respondents' gift card preferences were also noted in the database.

Three weeks after the priority mailer was sent out, a postcard was mailed to the urban non-respondent households alerting to them to the upcoming field effort. The field effort began one week later on October 30, 2009, and lasted six weeks. The in-person effort consisted of three attempts at the door made at different times of the day. Field visits took place between 8 a.m. and 8 p.m. After the third attempt, if the survey was not secured, the field researcher left a survey, a business reply envelope, and a letter.

For planning purposes, each field researcher was assigned to a region of the state a few weeks in advance. Each region was visited at least once. The amount of time allotted for each region depended on the number of surveys needed to reach the 50 percent completion target for the ZIP codes within the region. The schedule was updated on a weekly basis to reflect the progress made to date.

Each field researcher was assigned a phone scheduler for the week. Using the database to obtain the most up-to-date information, the phone scheduler focused on the ZIP codes that had the highest need first. To create a detailed weekly itinerary, the sites from the ZIP code were mapped and grouped by location. Then, a contact attempt was made. The scheduler notified the household that field personnel would be in their area and scheduled an appointment for the field staff when possible. Field researchers were expected to complete at least one ZIP code per day.

As a ZIP code approached the 50 percent complete target, it dropped into a *call-only* status. There was another group of phone schedulers who attempted to reach this population. During this round of phone calls, participants were asked to complete the survey and mail it in, complete it online, or complete it with the caller over the phone right then.

The field researchers received a daily schedule that included the participant's name, telephone number, and address; the order of which sites were to be visited; and the time of any scheduled appointments. Also included was the number of completed surveys needed to reach the 50 percent targeted response rate and any notes collected during previous visits and phone calls. When a survey was completed, the participant was given a $10 gift card. If no one was home, the researcher left a card with their name and personal phone number and indicated when they would return. If it was the final visit to the home, a survey and business reply envelope were left along with the card.

All field collection was done in the same manner as the regular RASS survey in that the participant was expected to complete the survey. The field researchers helped fill out the survey only when the respondent indicated they were not able to do it themselves. It was not completed as an audit or interview process to ensure the quality of the data for the non-response sample was the same as the rest of the RASS respondents. Roughly half of the field researchers spoke Spanish. All of the field researchers carried Spanish surveys with them and were prepared to offer a Spanish-language survey if requested.

At the conclusion of each day, the field researchers sent back an update of the day's activities in a site report. For each home that was visited that day, a field visit result, a field visit time, and a field visit comment were required. This information was uploaded into the database daily.

Non-Response Follow-Up Survey Completes

As shown in Table 2-7, of the 5,671 non-response surveys mailed, 2,323 (41 percent) surveys were returned: 246 from rural areas and 2,077 from urban areas. Of the 2,077 surveys returned from urban areas, 1,494 responded by mail (71 percent), 455 (22 percent) responded as the result of the field outreach effort, and 128 (6 percent) responded online.

Table 2-7 shows the distribution of survey returns by utility and by response mode.

Table 2-7: Non-Response Completes by Utility and Mode

Urban/Rural	Utility	Mail	Online	Field	Grand Total
Rural	LADWP	4			4
	PG&E	150	4		154
	SCE	79	1		80
	SDG&E	8			8
Rural Total		241	5		246
Urban	LADWP	291	23	89	403
	PG&E	343	22	113	478
	SCE	761	74	222	1057
	SDG&E	99	9	31	139
Urban Total		1494	128	455	2077
Grand Total		**1735**	**133**	**455**	**2323**

Source: *2010 California Residential Appliance Saturation Survey*

The field effort was prioritized to achieve a minimum of 50 percent response rates within ZIP code clusters. Table 2-8 presents the success of that effort. Sixty percent of the 262 ZIP code clusters had response rates over 40 percent, and only 12 (6 percent) clusters had response rates lower than 30 percent.

Table 2-8: Response Rates Within ZIP Code Clusters

ZIP code Cluster Returns (%)	Number of ZIP codes
0-9%	0
10-19%	1
20-29%	15
30-39%	88
040-49%	109
50-59%	44
60-69%	5
70-79%	0
80-89%	0
90-100%	0

Source: *2010 California Residential Appliance Saturation Survey*

Survey Weights

Survey weights were assigned to both the individually metered and master-metered survey results. The results were weighted to the initial populations submitted by the participating electric utilities and calculated separately for individually metered and master-metered households.

The individually metered results contained both the initial mailing completes and the non-response follow-up completes. Separate weights were calculated for each of these to account for the different sampling approaches. Basic weights were calculated for the master-metered results, equal to the ratio of the population count divided by the count of completed surveys per stratum.

Individually Metered Sample Weights

The approach for calculating the weights for the individually metered sample for the 2009 RASS followed the same approach as the weighting scheme used for the 2003 RASS. This approach weights the non-response follow-up sample less heavily by assuming the follow-up sample represents only the follow-up population, not the entire set of non-responders to the initial mailing. In effect, the responding sample represents only the people who would have responded to the initial mailings or to the follow-up effort.

The SDG&E oversample and the SCE oversample were included in the calculations. Since the SDG&E oversample followed the sample sampling design as the Energy Commission-sponsored sample, the SDG&E oversample was simply rolled into the sample of completed surveys. The SCE oversample used an additional stratification variable, the T24 zone, so weights were calculated at the more detailed level of stratification.

The equations for the initial mail sample stratum weights (w_1) and the follow-up sample stratum weights (w_2) are presented below.

$$w_1 = \frac{\frac{N}{n_1} \times \frac{n_1}{n_s}}{\frac{n_1}{n_s} + \frac{n_2}{n_f} \times \left(1 - \frac{n_1}{n_s}\right)} \quad \text{and} \quad w_2 = \frac{\frac{N}{n_2} \times \frac{n_2}{n_f} \times \left(1 - \frac{n_1}{n_s}\right)}{\frac{n_1}{n_s} + \frac{n_2}{n_f} \times \left(1 - \frac{n_1}{n_s}\right)}$$

Where:

N = population

n_1 = response to initial mail survey

n_2 = response to follow-up survey

n_s = initial mail sample (number of initial surveys mailed)

n_f = follow-up sample

for each stratum.

In general, the weights for each stratum followed the equations above; the only variation was in strata where there were completes from one response category but not from the other. For example, if there were surveys completed from the initial mailing category, but not any surveys completed from the follow-up category. In that case, the equation became the same as for basic weights—the population divided by the count in the sample.

Table 2-9 presents the completed surveys by response groups with sample weights by stratum for the individually metered sample. Sample weights for the initial mailing ranged from a low of 14.3 to a maximum of 1,821. Sample weights for the non-response sample ranged from 32 to 14,703.27. The columns of the table include the following information:

- Columns A through E indicate the stratification variables.
- Column F contains the prefix used for the SFCode in the database to indicate the stratum per respondent.
- Column G shows the population.
- Column H shows the actual number of completes from the initial mailings.
- Column I shows the actual number of completes from the non-response follow-up effort.
- Column J gives the sample weight for the initial mailing completes.
- Column K provides the sample weight for the non-response follow-up completes.

Table 2-9A: Individually Metered Weights (PG&E, LADWP, and SDG&E)

A	B	C	D	E	F	G	H	I	J	K
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	Population	Initial Mail Completes (sample 1)	Follow-Up Completes (Sample 2)	Weight 1	Weight 2
PG&E	OLD	NO	SF HIGH	1	P1	71105	138	3	389.422	5788.23
				2	P2	133312	189	21	248.355	4113
				3	P3	291702	397	41	254.428	4651.07
				4	P4	624767	935	76	277.335	4808.66
				5	P5	321631	439	40	281.496	4951.36
			SF LOW	1	P6	50944	104	7	234.837	3788.71
				2	P7	48826	76	6	244.742	5037.6
				3	P8	111792	174	10	301.691	5929.78
				4	P9	352079	572	42	291.637	4411.01
				5	P10	320259	526	55	240.96	3518.44
			MF	3	P11	93666	97	20	221.433	3609.35
				4	P12	279037	449	40	205.414	4670.15
				5	P13	383052	547	67	197.861	4101.82
				ALL Other	P14	61274	96	4	254.702	9205.65
		YES	SF HIGH	1	P15	90273	176	13	225.542	3890.59
				2	P16	32859	60	6	204.347	3433.03
				3	P17	70859	125	7	298.563	4791.23
				4	P18	53825	95	5	307.023	4931.56
				5	P19	24196	61	4	179.525	3311.24
			SF LOW	1	P20	37705	78	5	238.639	3818.23
				4	P21	24125	39	2	336.003	5510.45
				5	P22	19418	45	4	178.803	2842.97
				ALL Other	P23	23945	54	4	200.656	3277.39
			MF	4	P24	120724	170	22	178.068	4111.47
				5	P25	127945	203	20	201.202	4355.05
				ALL Other	P26	30913	57	4	200.647	4869.03
	NEW	NO	SF HIGH	2	P27	26846	36	7	184.297	2887.33
				3	P28	44729	48	7	295.566	4363.12
				4	P29	44330	47	7	286.37	4410.09
				ALL Other	P30	20156	55	3	198.163	3085.68
			SF LOW	ALL	P31	72593	106	13	233.606	3679.29
			MF	ALL	P32	72837	91	16	191.494	3463.19
		YES	ALL	ALL	P33	37242	40	3	421.29	6796.81
	Smart Meter	NO	SF HIGH	2	P34	31574	56	5	205.39	4014.44
				3	P35	123154	161	18	245.267	4648.11
				ALL Other	P36	15912	48	6	111.162	1762.71
			SF LOW	ALL	P37	69894	98	9	222.592	5342.22
			MF	ALL	P38	53876	41	7	289.107	6003.23
		YES	ALL	ALL	P39	16031	28	4	159.33	2892.44
LADWP	OLD	NO	L	11	L1	355912	549	82	214.384	2905.06
				12	L2	106221	152	40	170.918	2006.04
			M	11	L3	226484	388	50	192.688	3034.42
				12	L4	143352	234	42	198.074	2309.59

A	B	C	D	E	F	G	H	I	J	K
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	Population	Initial Mail Completes (sample 1)	Follow-Up Completes (Sample 2)	Weight 1	Weight 2
			H	11	L5	109446	213	34	184.824	2061.13
				12	L6	165036	327	66	175.844	1629.32
	NEW		L	11	L7	61028	106	18	153.521	2486.38
				12	L8	14003	34	18	56.984	670.31
			M	11	L9	32111	48	7	164.954	3456.17
				12	L10	15283	64	17	61.422	667.77
			H	11	L11	14133	46	13	76.919	814.98
				12	L12	16710	68	16	69.299	749.85
	ALL	YES	ALL	ALL	L13	10258	39	3	106.301	2037.43
SDG&E	OLD	NO	L	13	G1	319860	945	44	118.183	4731.3
			M	13	G2	423553	1387	55	107.541	4988.97
			H	13	G3	306876	970	30	119.73	6357.95
		YES	L	13	G4	6752	67	4	42.302	979.45
			M	13	G5	17784	64	2	130.218	4725.04
			H	13	G6	32414	90	2	195.265	7420.07
	NEW	NO	M	13	G7	24889	74	5	95.964	3557.53
			H	13	G8	22134	75	2	134.889	6008.68
		All Other	All Other	13	G9	19854	67	3	102.006	4339.87

Source: *2010 California Residential Appliance Saturation Survey*

Table 2-9B: Individually Metered Weights (SCE)

A	B	C	D	E	F.1	F.2	G	H	I	J	K
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Population	Initial Mail Completes (sample 1)	Follow-Up Completes (Sample 2)	Weight 1	Weight 2
SCE	OLD	NO	SF HIGH	7	S1	13	76,096	195	20	119	2644.52
						14	6,098	11	1	225.85	3613.63
						16	13,960	63	8	93.44	1009.16
				8	S2	6	200,455	374	47	208.2	2608.27
						8	236,921	393	75	193.78	2143.56
						9,10	95,905	140	25	210.52	2657.3
						5	3,128	42	3	25.81	681.35
				9	S3	6	25,165	34	3	264.89	5386.19
						8	111,618	135	40	220.57	2046.01
						9	321,692	516	76	197.58	2891.3
						14	54,874	99	14	165.47	2749.43
						16	6,932	23	3	131.62	1301.58
				10	S4	10	363,253	593	88	203.57	2756.1
						14	66,264	134	7	220.39	5247.39
						15	57,029	205	19	97.01	1954.86
						16	25,457	78	1	191.65	10508.6
				Other	S5	6	15,273	29	4	229.83	2152.01
						8	26,658	48	1	288.19	12824.66
						9	205	1	.	205	.
						10,14,15,16	219	2	.	109.5	.
						99	133	2	.	66.5	.
			SF LOW	7	S6	13	27,636	144	22	44.72	963.45

A	B	C	D	E	F.1	F.2	G	H	I	J	K
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Population	Initial Mail Completes (sample 1)	Follow-Up Completes (Sample 2)	Weight 1	Weight 2
[redacted]						14	2,828	5	1	125.69	2199.56
						16	11,492	66	7	77.82	907.95
				8	S7	6	154,672	301	26	204.65	3579.66
						8	147,455	280	29	229.01	2873.56
						9,10	43,676	77	13	204.34	2149.36
						5	3,408	55	4	29.76	442.74
				9	S8	6	28,462	50	4	241.2	4100.46
						8	123,458	153	24	254.11	3524.16
						9	163,335	283	29	234.85	3340.43
						14	20,361	39	5	123.03	3112.59
						16	4,120	24	4	49.52	732.88
				10	S9	10	123,315	198	26	208.73	3153.32
						14	31,669	74	8	109.2	2948.49
						15	17,764	197	11	39.55	906.66
						16	25,766	143	17	77.89	860.44
				Other	S10	6	8,616	26	1	178.02	3987.57
						8,9,10,14,16	8,618	37	3	101.39	1622.21
						99	93	2	.	46.5	.
			MF	7	S11	13	21,591	207	10	49.65	1131.26
						14	2,253	8	.	281.63	.
						16	10,928	130	10	43.56	526.49
				8	S12	6	145,111	261	38	172.5	2633.9
						8	152,598	205	43	178.39	2698.34
						9,10	38,593	65	14	181.1	1915.83
						5	1,623	18	1	37.74	943.6
				9	S13	6	52,837	63	9	163.56	4725.89
						8	75,470	98	24	201.89	2320.21
						9	110,480	147	22	181.64	3808.12
						14	15,615	56	6	94.52	1720.3
						16	5,621	45	3	50.89	1110.32
				10	S14	10	112,920	232	30	143.22	2656.45
						14	19,585	81	7	83.09	1836.35
						15	21,174	284	14	33.48	833.27
						16	7,370	37	5	47.81	1120.19
				Other	S15	6	7,708	24	2	94	2726
						8	7,158	22	2	115.45	2309.03
						14	73	1	.	73	.
						9,10,16	128	2	2	32	32
						99	38	1	.	38	.
		YES	SF HIGH	8	S16	6	11,280	39	4	123.18	1618.97
						8	9,339	25	1	177.32	4905.93
						9,10	5,591	9	.	621.22	.
						5	392	12	.	32.67	.
				9	S17	6	1,080	3	1	108	756
						8	2,648	12	3	73.56	588.44
						9	8,028	24	1	203.61	3141.39
						14	2,200	16	.	137.5	.
						16	4,626	26	2	66.56	1447.71
				10	S18	10	15,953	41	7	142.09	1446.75
						14	8,230	24	3	144.39	1588.25

A	B	C	D	E	F.1	F.2	G	H	I	J	K
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Population	Initial Mail Completes (sample 1)	Follow-Up Completes (Sample 2)	Weight 1	Weight 2
[redacted]						15	5,861	16	2	83.73	2260.67
						16	1,817	11	.	165.18	.
				Other	S19	6	1,335	7	2	52.01	485.45
						8	805	3	.	268.33	.
						9,10	42	1	.	42	.
						13	3,701	20	1	56.08	2579.48
						14	485	5	.	97	.
						16	3,027	27	4	43.33	464.26
						99	2,389	72	3	19.33	332.45
			SF LOW	ALL	S20	6	8,705	15	1	397.94	2735.86
						8	11,702	16	4	198.82	2130.22
						9	9,923	11	2	257.07	3547.6
						10	4,658	3	.	1552.67	.
						13	891	3	.	297	.
						14	4,230	17	.	248.82	.
						15	3,404	33	1	71.48	1045.32
						16	6,887	30	3	96.55	1330.2
						99	1,027	14	1	44.01	410.8
						5	228	3	1	30.4	136.8
			MF	8	S21	6	75,631	139	19	182.08	2648.49
						8	56,001	76	15	184.93	2796.44
						9	10,967	20	7	120.3	1223.01
						5	1,041	11	.	94.64	.
				9	S22	6	17,334	40	2	219.73	4272.46
						8	13,650	21	4	193.16	2398.41
						9	30,464	45	9	150.44	2632.69
						14	3,287	9	1	126.42	2149.19
						16	2,359	26	1	39.1	1342.41
				10	S23	10	33,355	42	7	222.82	3428.06
						14	1,720	12	2	54.6	532.38
						15	7,004	51	2	49.41	2242.02
						16	3,489	22	1	54.73	2284.95
				Other	S24	6	2,293	5	2	38.22	1050.96
						8	1,488	5	.	297.6	.
						13	2,281	23	4	25.57	423.25
						14	268	4	.	67	.
						9,16	727	13	4	21.66	111.37
						99	8,688	257	6	18.5	655.38
	NEW	NO	SF HIGH	8	S25	5,6	11,529	21	2	178.47	3890.59
						8	12,540	13	4	88.31	2847.99
						9,10	4,548	10	2	121.82	1664.89
				9	S26	6	738	2	.	369	.
						8	1,821	1	.	1821	.
						9	11,833	15	2	154.68	4756.4
						14	12,328	23	3	205.16	2536.47
						16	654	4	.	163.5	.
				10	S27	10	86,432	121	23	229.92	2548.32
						14	23,442	42	1	208.07	14703.27
						15	10,640	26	4	140.62	1745.99
						16	2,276	7	.	325.14	.

A	B	C	D	E	F.1	F.2	G	H	I	J	K
Electric Utility	Home Age	Electric Heat Presence	Home Type	CEC Forecast Climate Zone	Strata	T24	Population	Initial Mail Completes (sample 1)	Follow-Up Completes (Sample 2)	Weight 1	Weight 2
[redacted]				Other	S28	6,8	557	1	2	61.89	247.56
						13	13,361	38	3	114	3009.65
						14	1,637	6	1	154.43	710.4
						9,16	2,570	20	1	38.36	1802.84
			SF LOW	8	S29	6	9,565	11	2	109.94	4177.82
						8,9,10	16,554	21	4	177.15	3208.44
						5	52	1	.	52	.
				10	S30	10	30,916	40	7	215.96	3182.53
						14	10,951	28	.	391.11	.
						15	3,180	29	2	41.48	988.57
						16	1,478	6	.	246.33	.
				Other	S31	6	1,687	3	.	562.33	.
						8	2,459	2	.	1229.5	.
						9,10	5,681	16	.	355.06	.
						13	4,739	26	2	60.31	1585.41
						14	5,058	12	2	117.63	1823.23
						16	2,469	12	1	69.88	1630.47
			MF	Other	S32	5,6	7,529	16	3	143.14	1746.27
						8	7,750	19	2	140.91	2536.36
						9	7,478	24	5	81.98	1102.12
						10	9,186	12	3	95.19	2681.23
						13	1,290	20	.	64.5	.
						14	2,922	19	.	153.79	.
						15	1,429	17	.	84.06	.
						16	805	11	1	27.44	503.13
		YES	SF HIGH	All	S33	6	352	1	.	352	.
						8	312	1	.	312	.
						9	375	0	1	83.33	375
						10	1,360	7	.	194.29	.
						13	309	1	1	34.33	274.67
						14	383	0	2	57.45	191.5
						15	332	2	.	166	.
						16	217	3	.	72.33	.
			SF LOW			6	666	3	1	86.87	405.39
						8	607	1	1	37.94	569.06
						9	563	3	.	187.67	.
						10	1,339	5	1	93.42	871.91
						13	199	4	.	49.75	.
						14	235	1	.	235	.
						15	235	3	.	78.33	.
						16	130	2	.	65	.
			MF			5,6	2,082	8	1	109.58	1205.37
						8	1,351	4	.	337.75	.
						9	385	3	.	128.33	.
						10,16	2,560	6	1	177.43	1495.45
						13	329	4	1	14.3	271.78
						14	127	0	1	25.4	127
						15	282	3	.	94	.
Total							11,093,798	22141	2323		

Source: *2010 California Residential Appliance Saturation Survey*

Master-Metered Sample Weights

Basic weights were developed for the master-metered completed surveys as the ratio of the stratum population divided by the number of completed surveys. The population counts were from the initial population data as provided by the participating utilities.

Table 2-10 presents the number of completed surveys per stratum along with the sample weight for the master-metered sample.

Table 2-10: Master Metered Weights

A	B	C	D	E
Utility	Home Type	Population (Units)	Mail Completes	Weight
PGE	2-4 units	37,946	72	527.03
	Multi-family 5-20 units	17,699	35	505.69
	Multi-family >20 units	43,918	81	542.20
	Mobile home >4 units	105,112	229	459.00
SDGE	2-4 units	7,694	26	295.92
	Multi-family 5-20 units	5,558	13	427.54
	Multi-family >20 units	9,235	12	769.58
	Mobile home >4 units	33,469	144	232.42
SCE	2-4 units	10,789	49	220.18
	Multi-family 5-20 units	12,695	15	846.33
	Multi-family >20 units	19,974	28	713.36
	Mobile home >4 units	107,862	530	203.51
LADWP	2-4 units	1,480	3	493.33
	Multi-family 5-20 units	3,042	3	1014.00
	Multi-family >20 units	9,676	7	1382.29
	Mobile home >4 units	3,775	10	377.50
Total		429,924	1,257	

Source: *2010 California Residential Appliance Saturation Survey*

Comparison of Results Across Sampling and Study Groups

Non-Response Follow-Up Comparison

A non-response follow-up effort can effectively reach segments of the population that do not respond to the initial mailings. Table 2-11 presents a comparison of the households that completed their surveys in response to the initial mailings to the households responding to the non-response follow-up effort. The non-response households had similar major equipment and energy usage in their households to the initial mail responders. Key differences of non-response follow-up households were:

- Less likely to own their residence.
- Likely to have fewer seniors in the household.
- Less likely to use English as their primary language.
- More likely to have a head of household that is Hispanic.

Table 2-11: Comparison by Surveying Method

	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	13,968	1,389	3,599	412	3,758	480	816	42
Weighted to Population	2,716,013	4,333,328	562,229	1,243,344	589,620	1,443,735	103,337	102,191
Average Electric Consumption	7,549	7,611	4,226	4,127	3,559	3,744	5,540	5,560
Average Gas Consumption	427	418	240	236	155	147	334	345
Average Dwelling Size	1,911	1,864	1,203	1,131	955	927	1,277	1,353
Average Dwelling Age	37.8	37.0	34.6	34.6	31.9	32.0	28.1	28.2
Average Number of People	2.82	3.39	2.54	2.79	2.09	2.43	2.13	2.63
Average Number of Seniors	0.61	0.35	0.42	0.21	0.40	0.20	0.79	0.37
Average Income	79,062	80,001	58,253	56,341	50,859	55,686	32,970	46,373
Owners	91%	86%	49%	33%	28%	22%	86%	84%
Central Cooling	59%	60%	46%	41%	43%	42%	70%	73%
Gas Space Heating	83%	86%	77%	74%	60%	62%	62%	51%
All Exterior Walls Insulated	57%	56%	45%	41%	43%	45%	60%	53%
CFL Penetration	87%	84%	85%	83%	84%	80%	88%	74%
Primary Language English	91%	84%	82%	74%	85%	76%	94%	95%
Head of Household Hispanic	17%	27%	23%	32%	18%	26%	11%	17%
College Grad or Higher	56%	54%	50%	47%	53%	52%	22%	20%

Source: *2010 California Residential Appliance Saturation Survey*

Online Survey Comparison

An online survey option can reach respondents who are not inclined to complete a paper and pencil survey. Although only 3 percent of surveys returned were completed online, the option may have attracted some households that may not have returned a survey had the option been not available.

Table 2-12 presents a comparison of respondents by whether they completed a paper survey or submitted a survey through the website. The low number of online responses makes it difficult to draw strong conclusions. The results suggest that households completing a survey online were more likely to live in a newer dwelling, less likely to have seniors in the household, and more likely to have insulation in exterior walls. They were also more likely to have a higher income, have a college degree or higher, and use English as their primary language.

Table 2-12: Comparison by Response Method

	Single Family		Multi-Family (2-4 Units)		Multi-Family (5+ Units)		Mobile Homes	
	Mail Survey	Online Survey	Mail Survey	Online Survey	Mail Survey	Online Survey	Mail Survey	Online Survey
Completed Surveys	14,897	460	4,048	130	4,261	154	1,754	17
Weighted to Population	6,888,657	160,684	1,816,960	58,090	2,087,375	56,209	454,025	1,718
Average Electric Consumption	7,583	7,770	4,138	4,756	3,679	4,079	5,550	5,537
Average Gas Consumption	422	404	236	276	149	156	339	368
Average Dwelling Size	1,882	1,880	1,167	1,317	923	1,025	1,212	1,339
Average Dwelling Age	37.5	30.7	35.3	31.8	33.3	21.3	28.4	23.3
Average Number of People	3.17	3.02	2.72	2.50	2.31	1.83	2.11	2.29
Average Number of Seniors	0.46	0.18	0.30	0.08	0.28	0.04	0.72	0.56
Average Income	79,341	92,422	56,311	97,121	52,147	102,238	34,701	51,337
Owners	88%	85%	39%	57%	22%	28%	88%	97%
Central Cooling	59%	62%	42%	49%	41%	44%	66%	29%
Gas Space Heating	85%	86%	74%	82%	61%	57%	68%	48%
All Exterior Walls Insulated	56%	65%	41%	54%	44%	50%	55%	77%
CFL Penetration	85%	90%	83%	87%	82%	78%	84%	93%
Primary Language English	86%	90%	77%	86%	79%	88%	95%	92%
Head of Household Hispanic	24%	12%	28%	29%	23%	10%	13%	8%
College Grad or Higher	54%	67%	47%	70%	51%	91%	20%	24%

Source: *2010 California Residential Appliance Saturation Survey*

Master-Metered Comparison

The master-metered survey response rates were lowest in multifamily complexes of five or more units and highest from households in mobile-home parks. The overall low number of responses from master-metered households makes it difficult to make strong conclusions when compared to the much larger sample of individually metered households.

Table 2-13 provides a comparison of the responses received from individually metered households to those received from master-metered households. In general, households in the master-metered mobile-home sample have similar characteristics to the households in the individually metered sample, except for reporting a lower income. The multifamily complex households on master meters are more likely to live in older dwellings, have more seniors as members of their households, and use English as their primary language.

Table 2-13: Comparison of Individually Metered to Master-Metered Household Results

	Single Family	Multi-Family (2-4 Units)		Multi-Family (5+ Units)		Mobile Homes	
	Individual Metered	Individual Metered	Master Metered	Individual Metered	Master Metered	Individual Metered	Master Metered
Completed Surveys	15,357	4,011	167	4,238	177	858	913
Weighted to Population	7,049,341	1,805,573	69,477	2,033,355	110,230	205,529	250,215
Average Dwelling Size	1,882	1,154	1,719	935	704	1,315	1,117
Average Dwelling Age	37.3	34.6	50.9	32.0	52.3	28.1	28.6
Average Number of Seniors	0.45	0.28	0.63	0.26	0.52	0.58	0.84
Average Income	79,639	56,936	77,048	54,287	36,219	39,634	30,323
Owners	88%	38%	75%	23%	6%	85%	90%
Central Cooling	59%	43%	27%	42%	31%	71%	61%
Gas Space Heating	85%	75%	65%	61%	62%	56%	77%
All Exterior Walls Insulated	56%	42%	34%	45%	36%	56%	54%
Clothes Washer	85%	83%	86%	81%	82%	81%	86%
Primary Language English	86%	76%	90%	78%	90%	95%	95%
Head of Household Hispanic	23%	29%	20%	23%	14%	13%	13%
College Grad or Higher	55%	47%	59%	52%	51%	21%	19%

Source: *2010 California Residential Appliance Saturation Survey*

Energy Consumption Comparison

Household energy consumption was compared between RASS survey respondents and the target population. The energy consumption data were obtained from the original population files, from which the sample frame was developed. The energy consumption of the survey respondents was then compared to the average energy consumption of the population by stratum. Table 2-14 presents the comparison of the average energy consumption for respondents compared to the target population. The average energy consumption of respondents in the higher energy consumption strata was slightly lower than the population averages for their respective stratum. Conversely, the average energy consumption of respondents in the lower energy consumption strata was slightly higher than the average consumption for the population for their stratum.

The *All* stratum column includes households that were aggregated across strata because of low numbers in the more detailed strata. Because these strata reflect combinations of household types, the average for the respondents is more likely to vary more from the population compared to the more homogeneous strata.

Table 2-14: Comparison of Energy Consumption for RASS Respondents and Target Population

Usage By Household		Dwelling Type and Usage Strata Definitions							
Utility		High	Med	Low	SF-High	SF-Low	MF	All	Utility Totals
PG&E	Population Count				2,021,230	1,131,580	1,223,324	53,273	4,429,407
	Respondent kWh/Year				10,056	3,633	4,204	6,959	6747
	Population kWh/Year				10,337	3,484	4,136	7,086	6835
	Average Error				-2.7%	4.3%	1.6%	-1.8%	-1.3%
SCE	Population Count				1,996,710	1,079,835	1,126,426	17,327	4,220,298
	Respondent kWh/Year				10,059	3,613	5,118	5,623	6,855
	Population kWh/Year				10,428	3,525	4,493	6,032	7,107
	Average Error				-3.5%	2.5%	13.9%	-6.8%	-3.5%
SDG&E	Population Count	361,424	466,226	326,612				19,854	1,174,116
	Respondent kWh/Year	11,169	5,157	2,294				2,336	6,134
	Population kWh/Year	11,731	5,113	2,211				2,228	6,295
	Average Error	-4.8%	0.9%	3.8%				4.8%	-2.6%
LADWP	Population Count	305,325	417,230	537,164				10,258	1,269,977
	Respondent kWh/Year	12,914	5,314	2,286				5,545	6,412
	Population kWh/Year	13,195	5,183	2,213				6027	5,860
	Average Error	-2.1%	2.5%	3.3%				-8.0%	9.4%
Totals	Population Count	666,749	883,456	863,776	4,017,940	2,211,415	2,349,750	100,712	11,093,798
	Respondent kWh/Year	11,869	5,212	2,290	10,058	3,622	4,776	5,058	6,654
	Population kWh/Year	12,401	5,146	2,212	10,382	3,504	4,307	5,839	6,770
	Average Error	-4.3%	1.3%	3.5%	-3.1%	3.4%	10.9%	-13.4%	-1.7%

Source: *2010 California Residential Appliance Saturation Survey*

Precision of RASS Estimates

Individually Metered Sample Precision

Table 2-15 presents the precision of estimates for the individually metered sample by electric utility at the 90 percent confidence intervals. The three columns on the right of the table provide the percentage points to be added to and subtracted from an estimate of 50 or 50, 20 or 80, and 10 or 90 percent, respectively, to obtain the 90 percent confidence bounds.

Table 2-15: Precision of Estimates for the Individually Metered Sample

			90% Confidence Bounds (+/-)		
Utility	Population	Total Completes	50/50%	20/80%	10/90%
PG&E	4,429,407	7,390	1.0%	0.8%	0.6%
SCE	4,220,298	10,514	0.8%	0.6%	0.5%
SDG&E	1,174,116	3,886	1.3%	1.1%	0.8%
LADWP	1,269,977	2,674	1.6%	1.3%	1.0%
Total	11,093,798	24,464	0.5%	0.4%	0.3%

Source: *2010 California Residential Appliance Saturation Survey*

Master-Metered Sample Precision

Table 2-16 presents the precision of estimates for the master-metered sample by electric utility at the 90 percent confidence intervals. The three columns on the right of the table provide the percentage points to be added to and subtracted from and estimate of 50 or 50, 20 or 80, and 10 or 90 percent, respectively, to obtain the 90 percent confidence bounds.

Table 2-16: Precision of Estimates for the Master-Metered Sample

			90% Confidence Bounds (+/-)		
Utility	Population	Total Completes	50/50%	20/80%	10/90%
PG&E	204,675	417	4.0%	3.2%	2.4%
SCE	151,318	622	3.3%	2.6%	2.0%
SDG&E	55,955	195	5.9%	4.7%	3.5%
LADWP	17,973	23	17.2%	13.7%	10.3%
Total	429,921	1,257	2.3%	1.9%	1.4%

Source: *2010 California Residential Appliance Saturation Survey*

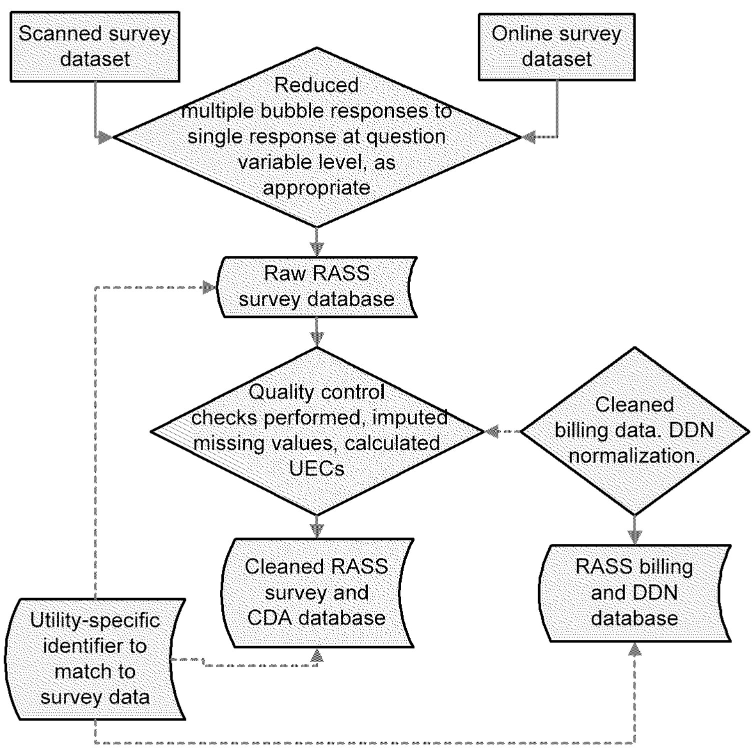
CHAPTER 3: Database Preparation

This section describes the three databases that were delivered to the Energy Commission and each of the participating electric and gas utilities upon the completion of the study. The three databases from the RASS study included:

- *Raw RASS survey database*: This database contained RASS survey data that were subjected to minimal cleaning procedures, i.e., limited changes to the responses that were marked on the survey.
- *Cleaned RASS survey database and conditional demand analysis (CDA) database*: This database contained RASS survey data, variables used in the CDA and household and end-use unit end consumption (UEC) estimates.
- *Billing and Degree-Day Normalization (DDN) database*: This contained cleaned billing series data from each of the participating utilities and estimated normalized annual consumption for each household.

The participating utilities received an additional database, which contained utility-specific information, allowing them to match the RASS survey and billing data to their specific customers. Figure 3-1 provides an overview of how these databases were constructed.

Figure 3-1: Overview of RASS Database Preparation



Source: *2010 California Residential Appliance Saturation Survey*

Database 1: Raw RASS Survey Database

The majority of RASS surveys were completed as paper surveys and scanned electronically into a fixed-format text file. Data from surveys completed online were converted to the same fixed-format file structure as the scanned paper survey data file. Responses from the paper surveys and the online surveys were then combined into one dataset.

The survey had a total of 1,416 potential responses to questions, each represented by a bubble that was recorded in a text file by the scanning program. The initial SAS code created separate fields for each response bubble as the text files were read into SAS datasets.

The first data cleaning step entailed condensing each of the separate fields into a single variable by assigning a value based on the populated bubble. For questions in which a respondent marked one response, the variable was simply assigned the value of the single response. For cases where a respondent marked multiple responses, the study team developed a set of decision rules to select a single value to be assigned to each variable. The choice was typically programmed as either the minimum or maximum value of the multiple responses, depending on the specific variable. For example, for the variable indicating years of residence, the maximum value was chosen for respondents who had provided multiple answers. For some survey variables, the choice of the single value assigned from multiple values depended on responses to other questions within the survey, thereby providing logically consistent answers to each question. The variables contained in the Raw RASS Survey database are listed in the Appendices Volume, Appendix L.

Database 2: Cleaned RASS Survey and CDA Database

The study team conducted multiple quality control checks and performed some additional cleaning steps on the raw survey data to develop the cleaned RASS survey and CDA database. These steps resulted in the omission of surveys based on incomplete data, an inordinate number of multiple responses, or an excessive number of logical inconsistencies from the final dataset. The details of the survey cleaning processes are discussed in Chapter 4.

Algorithms were designed to fill and impute missing values for variables used in the CDA. In addition, responses indicating fuels used for space heating, water heating, and other appliances were cross-referenced with billing data to identify and correct fuel misreporting. The CDA data imputation and consumption cleaning processes are discussed in Chapter 4. The specifics of the CDA modeling process are covered in Chapter 5.

Household and end-use UECs from the CDA, and post-normalized annual household electric and gas consumption data were appended to the cleaned survey data. The variables contained in the Cleaned RASS Survey and CDA database are listed in the Appendices Volume, Appendix M.

Non Response Indicator

Some sets of questions in the RASS survey incorporated a skip pattern. For example, if a household did not pay for the energy use of a swimming pool, they were instructed to skip to the next survey section. The non-responses for questions subject to the intentional skip pattern were assigned a value of 99 as being not applicable during the cleaning process. The simple respondent non-response was assigned a value of 97 as a missing value during the cleaning process. Surveys that contained an excessive amount of non-responses were omitted from the cleaned survey dataset.

Logical Response Inconsistencies

Some survey questions were interrelated, to which the response to one question would presumably influence the response of another question. For example, if a household reported not having a gas line to the residence, it would be logically inconsistent if they reported having a gas range in the residence. Where possible, logically inconsistent responses were corrected using billing data or other survey information. In cases where a value could not be inferred, the response was assigned the missing value of 97 and a logical inconsistency flag was set. The number of logical inconsistency flags was counted. No surveys were omitted from the cleaned survey dataset from having too many logical inconsistencies.

Imputing Missing Values

Although missing survey values were recorded as 97s in the cleaned RASS survey database, retaining these missing values in the CDA would have resulted in a non-response bias. Therefore, an approach was developed to impute all of the variables that were used in the CDA. The approach used to impute these variables is discussed in Chapter 4.

Refining Fuel and System Types

Previous CDA studies conducted on the California residential population have shown that the misreporting of fuels used for space heating and water heating was common. Since space and water heating account for large shares of household energy consumption, the variables used in the CDA needed to accurately reflect the fuel type in the household for the results of the CDA to be accurate. The approach to fuel checking and imputing values is discussed in detail in chapter 4.

Estimated UECs

The household and end-use UECs from the CDA were appended to the cleaned survey data. Pre- and post-normalized annual household electric and gas consumption variables were also added to the database. In addition, the CDA required the normal heating and cooling degree-day series from the 2003 RASS to construct pre-CDA engineering estimates, based on the 2003 UEC equations. These temperature and daylight series were also included in the CDA database.

Database 3: Billing and Degree-Day Normalization Data File Set

The study team conducted quality control checks on the electric and gas consumption data prior to performing the degree-day normalization (DDN). Since the weather-normalized annual usage was calculated independently for electric and gas consumption, the data were stored in two separate files. The DDN is discussed in detail in Chapter 4. The DDN files contained read dates, number of days in billing cycle, usage for up to 21 electric or 20 gas billing cycles, and the pre-normalized annual consumption. The DDN process used daily average weather and normal weather mapped to respondents by the T24 zone, so these variables were also included in the files. The variables contained in the Billing Data and Degree Day Normalization file set are listed in the Appendices Volume, Appendix N.

Database Formats

The volume of data generated by the RASS study demanded the use of software with the capacity to manipulate large datasets and the ability to support the analyses required by the study. The study team used the Statistical Analysis System (SAS) software package from the SAS Institute to analyze the RASS data. All of the survey data, billing data, and weather data were stored as SAS datasets and analyzed from within the SAS environment. The description of the SAS files and code is contained in the Appendices Volume, Appendix O.

The final databases were provided in two file formats: a SAS dataset format and a comma delimited (.csv) format. The .csv format facilitated importing the data into other software packages.

The study team also updated the web interface from the 2003 RASS study that allows users to design their own queries to create reports directly from the RASS survey data. The website enhancements included providing the ability to compare RASS results between the 2009 and 2003 datasets and facilitating queries based on grouping by multiple variables.

Data Delivery

The RASS data were delivered to the Energy Commission and participating utilities on CD. The CD contained the files listed below.

- Raw survey data files
  - Individually metered: Min_max_output_new (.sas7dat, .csv) and min_max_output_new_CONTENTS (.xls)
  - Master-metered: MM_ Min_max_output_new (.sas7dat, .csv) and mm_min_max_output_new_CONTENTS (.xls)

- Cleaned survey and CDA data files
  - Survdata – unformatted (.sas7dat, .csv)
  - Survdatf – formatted (.csv)
  - SurvCONTENTS – file contents (.xls)
  - Formats – format statements (.txt)
  - ApplyFormats – applies formats to specific variables (.txt)
- Billing and degree-day normalization data files
  - DDN_electricbillingdatamodels (.sas7dat, .csv)
  - DDN_ electricbillingdatamodels_CONTENTS - file contents (.xls)
  - DDN_gasbillingdatamodels (.sas7dat, .csv)
  - DDN_ gasbillingdatamodels_CONTENTS - file contents (.xls)

CHAPTER 4: Data Cleaning And Processing

Overview

The section outlines the processing steps applied to the survey data to ensure the data used to develop estimates were as accurate as possible. These steps included eliminating surveys that were determined to have excessive amounts of invalid data, cleaning RASS Survey variables, and creating new variables through the cleaning process and the combination of survey variables. Figure 1 provides an overview of the general data cleaning process.

Paper Survey Processing

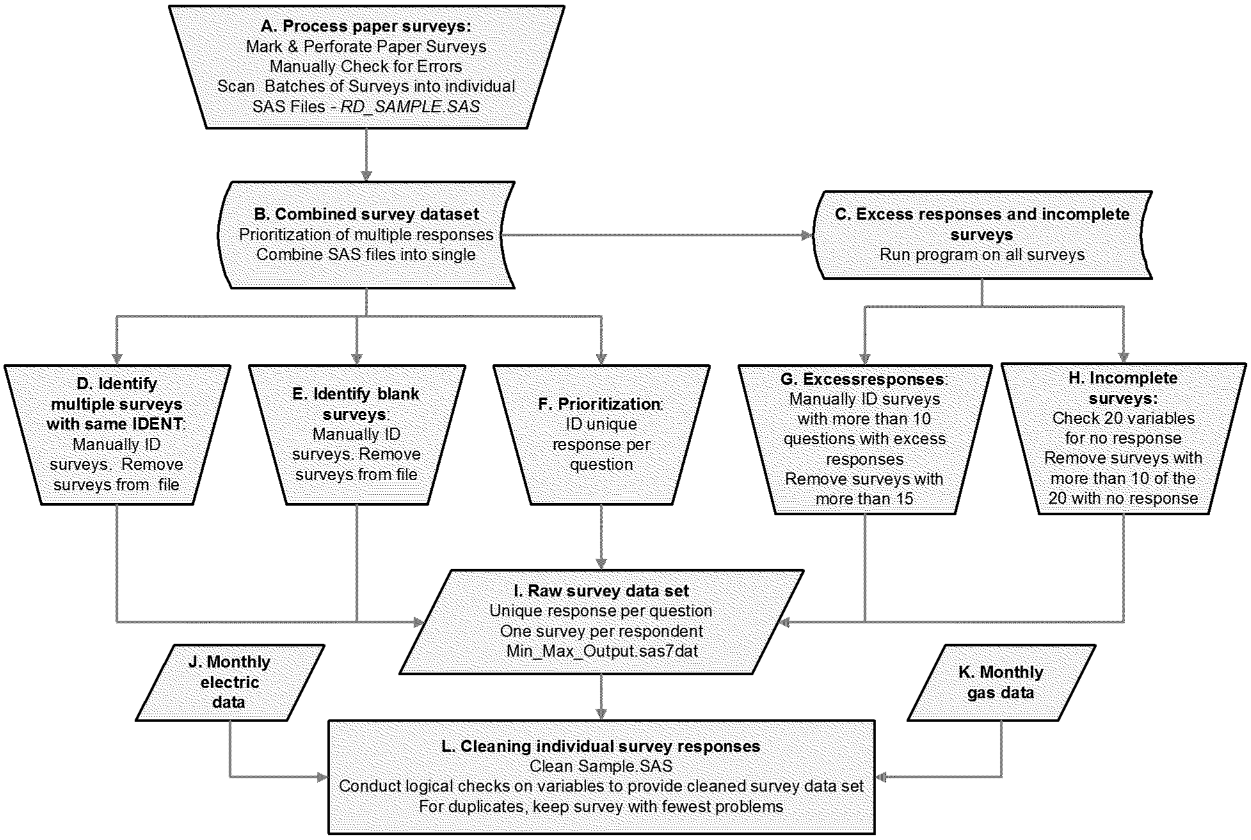
The first step in the survey cleaning process was to manually review and mark all survey responses with a marker to improve the accuracy of the scanning process. (See A of Figure 4-1.) Once this step was complete, batches of surveys were scanned into SAS data files such that each bubble response was reflected as a distinct variable.

Combined Survey Dataset

As shown in C of Figure 4-1, the SAS program *min_max.sas* created a single SAS dataset, *Alldata.sas7dat* that contained all scanned files, and also performed the following functions:

- *Identified duplicate surveys* – Multiple surveys with the same IDENT were identified (D of Figure 4-1). The paper surveys were reviewed to determine whether an error occurred in recording the barcode that contained IDENT during the scanning process. If the barcode was confirmed to be incorrectly recorded, the IDENT was corrected in the survey database. If the barcode had been recorded correctly, then the multiple surveys were considered duplicate surveys from the same household and were carried through the cleaning process to be resolved later by keeping the one with the fewest problems.
- *Identified blank surveys* –Surveys with responses missing for all questions were identified. The paper surveys were reviewed to determine whether they would be re-scanned or deleted (E of Figure 4-1).
- *Prioritization of multiple responses for each question* (F of Figure 4-1) – Some questions required a single answer, but the respondent provided multiple answers. For most questions, a unique response was inferred based on a set of predefined criteria for each question that picked either the minimum or maximum response category for that question. For certain survey questions, however, the mean response was used in place of the minimum or maximum response category. This process resulted in the SAS data set Min_MAX_Output.sas7dat.

Figure 4-1: Data Cleaning Process Overview



Source: *2010 California Residential Appliance Saturation Survey*

Excess Responses (G of Figure 4-1)

The program *TooManyResponses.SAS* was run on the initial survey dataset, *All_data.sas7dat*, to count the number of questions with excess responses (i.e. more responses were given than the question requested). Surveys containing more than 10 survey questions with excess responses were manually reviewed for scanning errors. If no errors in scanning were found, surveys with fifteen or more excess responses were flagged to be deleted. The number of questions with excess responses was recorded in the *Min_Max_output.sas7dat* dataset, and later used to assist in eliminating duplicate surveys. The process used to eliminate duplicate surveys is described in the section that describes cleaning individual survey questions.

Incomplete Surveys (H of Figure 4-1)

The *Too_Many_Responses.SAS* code also contained the five step process that identified incomplete surveys. This section outlines that process.

The first step in identifying incomplete surveys was to check a set of 20 variables for missing values. The variables were selected to represent the beginning, middle, and end of the survey.

The variables were divided into two groups, based on whether they represented a single question on the survey or a group of responses.

The first group contained 15 variables that represented individual survey questions that all respondents should have answered and that were not subject to a skip pattern.

Table 4-1: Survey Variables Used to Identify Incomplete Surveys

Survey Section	Survey Variable	Description
A	DWLTYPE	Dwelling type
	OWNRENT	Own or rent
	YRS_RES	Years of residence
	BUILTYR	Year home built
	NUMROOM	Number of rooms
	SQFT	Square footage
B	PAYHEAT	Pays for heat
C	PAYCOOL	Pays for cooling
D	PAYWH	Pays for hot water
E	LNDRYEQP	Laundry equipment in home
G	RFNUM	Number of refrigerators
I	SPTYP	Spa type
K	WORKHOME	Person works at home
M	WLWTRPMP	Well water pump
N	INCOME	Household income

Source: *2010 California Residential Appliance Saturation Survey*

The second group of variables used to identify incomplete surveys included five composite indicator variables. Shown in Table 4-2, the five composite indicator variables were each based on a set of survey variables that represented either multiple related survey questions or multiple sub-categories within a question. An example of the latter is the composite variable used to indicate missing values for both of the subcategories from the question on number of home computers, where the survey asks separately for number of laptops and the number of desktops.

Each composite variable considered the joint responses to the set of variables defining the composite indicator variable. If a respondent had missing values for each of the individual variables in the composite group, the composite indicator variable was coded as one. If at least one of the individual variables that made up the composite variable contained a non-missing value, then the composite variable was coded as zero to indicate there was information for at least one variable in the group.

Similar logic was used to construct composite variables for the presence of natural gas service, usage of various cooking appliances, number of exterior lighting fixtures, and presence of miscellaneous appliances.

Table 4-2: Composite Indicator Variables Used to Identify Incomplete Surveys

Survey Section	Composite Variable	Variables Needed	
		Survey Variables	Description
A	NGMISS	NGSERV	natural gas service in area
		NGLINE	natural gas line to house
F	COOKMISS	WRNUSE	Weekly ovenrange stovetop use
		WOVUSE	Weekly oven use
		WMWUSE	Weekly microwave oven use
K	NPCSMISS	NDSKPCS	number desktop PCs
		NLASPPCS	number laptop PCs
L	EXLIGHTMISS	EXINC	Number of exterior incandescent fixtures
		EXCFL	Number of exterior CFL fixtures
		EXLOWWV	Number of exterior low voltage light systems
		EXHID	Number of exterior HID fixtures
M	M1MISS	CHRGRS	Number of plug-in chargers
		FNPORT	Number of portable fans
		FNCEIL	Number of
		WNDATV	Number of wind turbine attic ventilators
		FNATTIC	Number of attic fans
		FNWHOLE	Number of ceiling fans
		AIRCLEAN	Number of wholehouse air cleaners
		HUMDEH	Number of humidifiers or dehumidifiers
		WINCLR	Number of wine or beverage coolers
		WHPURIFY	Number of water purification systems
		DHWRPMP	Number of domestic hot water recirculating pumps
		ELBLNKET	Number of electric blankets
		AQUAR	Number of aquariums
		TRSHCOMP	Number of trash compactors
		SAUNA	Number of saunas
		SCRTYSYS	Number of security systems
		POND	Number of pond or garden pumps
		GRGDROPN	Number of garage door openers
		LAWNMOWR	Number of electric lawn mowers

Source: *2010 California Residential Appliance Saturation Survey*

The second and third steps in identifying incomplete surveys considered whether large portions of the survey were left blank. Responses to 18 questions (A1 – A18) from the Home and Lifestyle questions were checked to determine whether all questions were skipped. Any survey in which this entire section was left blank was flagged. Similarly, responses to the Laundry, Food Preparation, and Refrigerator sections (questions E1 – G2) were checked to determine if all responses were missing, and surveys in which all responses were missing were flagged.

All surveys flagged in steps one through three were identified and the paper surveys were reviewed to confirm the survey had been scanned correctly. Surveys that were scanned incorrectly were rescanned.

The final step of this process was to select surveys for deletion if 10 or more of the variables from the two groups were missing or both sections reviewed in steps three and four were entirely missing. The number of questions with missing responses was also recorded in the *Min_Max_output.sas7dat* dataset to assist in eliminating duplicate surveys from the same household once the cleaning of individual survey questions was complete.

Cleaning Individual Survey Questions

Figure 4-1 (L) shows that the SAS program *Clean_Sample.sas* was used next to combine the unique survey responses (I of Figure 4-1) with monthly electric (J of Figure 4-1) and gas billing data (K of Figure 4-1). This program was used to clean the individual survey questions, which consisted of the following steps, all of which are described in detail within this chapter:

- Refined Fuel System Types: The survey data set was combined with monthly electric and gas billing data to identify households in which fuel used for heating and water heat was misreported.
- Identified Year-Round Residents: While the 2003 CDA attempted to account for partial year effects of seasonal residents and vacationers, the current CDA restricted the analysis to year-round residents. Monthly electric data were used to identify partial year residents and remove them from the CDA and saturation estimates.
- Coded Non-Response and Not Applicable Response: The cleaning process distinguished between non-responses that resulted from the intentional skip pattern in the survey and questions in which the response was left blank. The former were coded as 99, meaning not applicable, while truly missing responses were coded as 97.
- Determined Logical Response Inconsistencies: Many survey questions were interrelated, requiring responses to be logically consistent. For example, if a respondent indicated that they did not have natural gas service in their area, it would not be consistent for them to have a gas line to the residence. Where possible, logically inconsistent responses were corrected using information contained in billing data or in other survey responses. In cases where a value could not be inferred, the response was set to 97 to reflect a missing value.

  The number of logical inconsistencies in each survey was counted using a cumulative flag that added 1 for each occurrence. The number of inconsistencies was used to identify surveys that contained too many errors to include in the CDA and also to assist in eliminating duplicate surveys.

- Removed Duplicate Surveys: The initial processing of surveys revealed 352 individually metered and 164 mastered-metered duplicate surveys in the scanned file. These were identified as multiple surveys with the same IDENT. Once the paper copies were reviewed to confirm that they were true duplicates, each was carried through the program, Too_Many_Responses.sas code to count the number of multiple and missing responses. These surveys were then carried through the Clean_Sample.SAS code to identify the number of logical inconsistencies. Once both of these processes were complete, a variable called PROBLEMS was created that equaled the sum of each of these three measures. Duplicate surveys were then eliminated by keeping the version with the lowest value of the PROBLEMS variable. For duplicate records with equal values of the PROBLEMS variable, the first record in the file was retained.

- Imputed Missing Values: Although missing survey values were recorded as 97 in the cleaned survey data set, retaining these as missing values in the CDA would result in non-response bias. Therefore, an approach was developed to impute missing values for all variables used in the CDA.

Invalid Surveys

Based on the criteria identified above, surveys flagged due to an excessive number of excess, logically inconsistent, or missing responses were removed from the survey dataset used in the CDA and saturation tables.

Table 4-3 presents the number of surveys removed from the dataset according to the reasoning discussed above. In addition, 58 blank surveys were identified and removed during the initial processing of survey files.

Table 4-3: Summary of Invalid Surveys

Reasons for Eliminating Survey	Number Eliminated
Too many multiple responses	3
Incomplete survey	179
Too many logical inconsistencies	0

Source: *2010 California Residential Appliance Saturation Survey*

Survey Specific Cleaning

This section describes the logic used to identify illogical responses and clean individual survey questions. The section is organized in the same order as the survey, presented in Appendices Volume, Appendix A, which is divided into the following sections:

- Your Home and Lifestyle
- Space Heating
- Space Cooling
- Water Heating
- Laundry
- Food Preparation
- Refrigerators
- Freezers
- Spas and Hot Tubs
- Pools
- Entertainment and Technology

- Lighting
- Miscellaneous Appliances
- Household Information

Your Home and Lifestyle

The *Your Home and Lifestyle* section of the survey contained 21 questions, many of which are critical for other data cleanings and the CDA estimates. The process used to clean these variables is discussed below. Cleaning procedures used for some variables required cross references with other survey variables or billing data, or both. Cross references are clearly delineated below.

Type of Dwelling

The process used to clean the dwelling type variable (A1-*DWLTYPE*) is presented in detail because this variable is used extensively to estimate imputed values for other survey variables and serves as a key explanatory variable in the CDA process.

The original survey response values for the "type of building" (A1 – DWLTYPE) included the following:

- 1 is a single-family detached house
- 2 is a townhouse
- 3 is a 2-4 unit apartment or condominium
- 4 is a 5+ unit apartment or condominium
- 5 is a mobile home
- 6 is other

Cleaning of the DWLTYPE variable addressed missing, inconsistent, and ambiguous responses. First, survey respondents that did not provide an answer to this question were coded 97, to reflect a missing value. Second, DWLTYPE was checked against several other survey questions to see if they contradicted each other. Third, attempts were made to match respondents who answered 6 (Other) to the DWLTYPE question to a less ambiguous response category.

Individually Metered Surveys

The variable *RESIDENCE* was created to reflect each household's corrected dwelling type. If there was no problem with the original *DWLTYPE* response, the original value for *DWLTYPE* was retained as *RESIDENCE*. The process of creating the *RESIDENCE* variable used the following information:

- Survey responses to DWLTYPE, payment of heating, cooling, water heating, laundry systems, and square footage.
- Residence type code provided by each participating electric utility for the sample frame dataset (sixth digit of IDENT for individual metered households).
- Household's service street address.

The sixth digit of *IDENT* was used to create the variable *RESTYPE*, which reflected the utility codes for individual metered home type. The RESTYPE variable included the following:

- 1 or 2 are single-family dwellings,
- 3 is a multifamily dwelling,
- 4 is a low electric consumption dwelling,
- 5 is a medium electric consumption dwelling,
- 6 is a high electric consumption dwelling, and
- 0 is unknown.

The *RESIDENCE* variable for individually metered households was defined according to the rules outlined below.

- If DWLTYPE was equal to 2, 3, or 4 and the utility's RESTYPE code was equal to 1, 2, 4, 5, 6, RESIDENCE equaled the individual's response for DWLTYPE. In this situation the survey response overrides the utility's RESTYPE code (codes 1 and 2 were single family; 4, 5, and 6 were consumption codes).
- If DWLTYPE equaled to 1 and the utility's RESTYPE equaled to 3 (utility code for multifamily), proceeded through the following checks:
  - Reviewed the service address. If address ended in a number 1-4 or the letter A, B, C, or D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - Reviewed the service address. If the service address did not end in a letter or a number, checked if the respondent paid for a major system and if the survey response to square footage was less than 2,500.
  - If both checks were satisfied, then set RESIDENCE to 2.
  - If none of the above conditions was met, set RESIDENCE to 1.
- If DWLTYPE was equal to 6 and the utility's RESTYPE was equal to 3, proceeded through the following checks:
  - Reviewed the service address. If the address ended in a number 1-4 or a letter A-D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - If the address did not end in a number or a letter, set RESIDENCE to 2.
- If DWLTYPE was equal to 6 and RESTYPE was equal to zero (the utility did not know the RESTYPE), proceeded through the following checks:

- Reviewed the service address. If the address ended in a number 1-4 or a letter A-D, set RESIDENCE to 3.
- Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
- If the address did not end in a number or a letter and the survey response to square footage was greater than or equal to 2,500, set RESIDENCE to 1.
- If the address did not end in a number or a letter and the survey response to square footage was less than 2,500, set RESIDENCE to 2.

- If DWLTYPE was equal to 6 and RESTYPE was equal to 4, 5, or 6 (utility codes were based on consumption, not a dwelling type indicator), proceeded through the following checks.
  - Reviewed the service address. If the address ended in a number 1-4 or a letter A-D set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D set RESIDENCE to 4.
  - If the address did not end in a number of a letter, set RESIDENCE to 1.
- If DWLTYPE was equal to 6 and RESTYPE was equal to 1 or 2 (utility codes for single family), proceeded through the following checks.
  - Reviewed the service address. If the address ended in a number 1-4 or a letter A-D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - If the address did not end in a number or a letter, set RESIDENCE to 1.
- If DWLTYPE equaled 97 and RESTYPE equaled 1 or 2 (utility codes for single family), proceeded through the following checks.
  - Reviewed the service address. If the address ended in a number 1-4 or letter A-D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - If the address did not end in a number or a letter, set RESIDENCE to 1.
- If DWLTYPE equaled to 97 and RESTYPE was equal to 3 (utility code for multifamily), proceeded through the following checks:
  - Reviewed the service address. If the address ended in a number 1-4 or a letter A-D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - If the address did not end in a number of a letter, set RESIDENCE to 2.

- If DWLTYPE equaled 97 and RESTYPE equaled zero (utility code for unknown), proceeded through the following checks:
  - Reviewed the service address. If the address ended in a number 1-4 or a letter A-D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - If the address did not end in a number or a letter and the respondent did not pay for any major system, set RESIDENCE to 1.
  - If the address did not end in a number or a letter, the survey response to square footage was less than 2,500, and the respondent paid for at least one of the major systems, set RESIDENCE to 2.
- If DWLTYPE equaled 97 and RESTYPE equaled 4, 5, or 6 (utility code for consumption strata), proceeded through the following checks:
  - Reviewed the service address. If the address ended in a number 1-4 or a letter A-D, set RESIDENCE to 3.
  - Reviewed the service address. If the address ended in a number larger than 4 or a letter later than D, set RESIDENCE to 4.
  - If the address did not end in a number or a letter and the respondent did not pay for any major systems, set RESIDENCE to 1.
  - If the address did not end in a number or a letter and the respondent paid for at least one of the major systems, set RESIDENCE to 2.

This process resulted in *RESIDENCE* values of 1, 2, 3, 4 or 5 for all households. Out of 24,647 surveys processed, 2,134 contained a *RESIDENCE* different from what was reported by the respondent in DWLTYPE. For these 2,134 households, 1,927 were changed from *DWLTYPE*=Other (6) or Not Answered (97 or missing). Table 4-4 below presents all such changes by *DWLTYPE*, reason for change, *RESIDENCE* and *RESTYPE*.

Table 4-4 shows that *DWLTYPE* was missing by 1,669 survey respondents. From this total, 1,334 households were classified as Single Family Detached houses. Of these 1,334 survey respondents, 1,010 were assigned to Single Family Homes because the street address does not end in a number or letter. The remaining 324 dwellings were assigned to single family because the survey indicated that the major systems are not paid for by the landlord. Another 335 dwellings were reassigned to Apartments, Condominiums, Townhouses, or Duplexes from missing.

Table 4-4: Dwelling Type Cleaning Results for Individually Metered Households

DWLTYPE	Reason for Change	RESIDENCE	RESTYPE							Total
			Unknown	Single Family High	Single Family Low	Multi-family	Low Usage	Med Usage	High Usage	
Single Family Detached House	Service Address Ends with a Number or a Letter	Apartment or Condominium, 2-4 units	0	0	0	71	0	0	0	71
		Apartment or Condominium, 5+ units	0	0	0	113	0	0	0	113
	Landlord pays for at least one major system and SQFT =<2,500	Townhouse, Duplex or Rowhouse	0	0	0	23	0	0	0	23
Other	Service Address Ends with a Number or a Letter	Apartment or Condominium, 2-4 units	1	1	3	21		2	0	28
		Apartment or Condominium, 5+ units	2	4	7	42	5	4	1	65
	Address End NOT a Number or a Letter	Townhouse, Duplex or Rowhouse	0	0	0	36	0	0	0	36
	Landlord pays for at least one major system and SQFT =< 2,500	Townhouse, Duplex or Rowhouse	7	0	0	0	0	0	0	7
	Address End NOT a Number or a Letter	Single Family Detached House	0	37	47	0	18	13	7	122
No Response	Service Address Ends with a Number or a Letter	Apartment or Condominium, 2-4 units	4	4	7	42	3	4	1	65
		Apartment or Condominium, 5+ units	6	21	21	70	22	11	1	152
	Address End NOT a Number or a Letter	Single Family Detached House	0	607	403	0	0	0	0	1010
		Townhouse, Duplex or Rowhouse	0	0	0	85	0	0	0	85
	Landlord does not pay for any major system	Single Family Detached House	86	0	0	0	69	76	93	324
		Townhouse, Duplex or Rowhouse	9	0	0	0	13	7	4	33
Total			115	674	488	503	130	117	107	2,134

Source: *2010 California Residential Appliance Saturation Survey*

Master-Metered Surveys

The field *RESIDENCE*, the updated dwelling type of the master-metered respondents, was created similarly to the individually metered respondents. *DWLTYPE* recorded the responses of the master-metered survey participants and *RESTYPE* was provided by the utilities. *RESTYPE* for the master-metered units contained values 1, 2, 3, and 4. The descriptions of the values were:

- 1 was a 2- to 4-unit duplex, triplex or quadplex
- 2 was a 5-20 unit multifamily dwelling
- 3 was a 20 + unit multifamily dwelling
- 4 was a mobile home park

In some cases, *DWLTYPE* contained missing values, or the value *OTHER*. Inconsistencies were found between *DWLTYPE* and *RESTYPE* in some other cases. For all these cases, it was assumed that the information provided by the utility, contained in field *RESTYPE,* was correct.

In particular, the following types of inconsistencies were found between DWLTYPE and RESTYPE:

- DWLTYPE=1 and RESTYPE=1, 2, 3 or 4
- DWLTYPE=2 and RESTYPE=2, 3 or 4
- DWLTYPE=3 or 4 and RESTYPE=1 or 4
- DWLTYPE=5 and RESTYPE=1, 2 or 3

The cleaning code for master-metered respondents assumed that the values stored in *RESTYPE* were correct. Since *DWLTYPE* and *RESTYPE* could not be mapped perfectly, the following rules were adopted –

- If RESTYPE=1 then RESIDENCE=2.
- If RESTYPE=2 or 3, then RESIDENCE=4
- If RESTYPE=4 then RESIDENCE=5

Table 4-5 provides the counts of *DWLTYPE* by *RESIDENCE*, for all surveys where *DWLTYPE* differed from *RESIDENCE*. In particular, there were 269 master-metered surveys where *DWLTYPE* differed from *RESIDENCE*. For example, there were 92 surveys that reported *DWLTYPE*="Single Family" that were changed in *RESIDENCE* to townhouse, duplex, or row houses. Similarly, there were 22 surveys where *RESIDENCE* was updated from single family to mobile homes.

Table 4-5: Dwelling Type Cleaning Results for Master-Metered Households

DWLTYPE	Residence		
	Townhouse, duplex or row house	Apartments 5-20 units	Mobile Homes
	RESTYPE=1	RESTYPE= 2 or 3	RESTYPE=4
Single Family	92	11	22
Townhouse, duplex, or Row house	0	2	0
Apartment or Condonimium 2-4 units	20	0	1
Apartment or Condonimium 5 or more units	5	0	1
Mobile Home	3	1	0
Other	6	10	30
Not Answered	16	26	23
Total	142	50	77

Source: *2010 California Residential Appliance Saturation Survey*

When no inconsistencies between *DWLTYPE* and *RESTYPE* were found, the survey response of *DWLTYPE* was carried over to the *RESIDENCE* variable. This process resulted in imputed values for the RESIDENCE variable for all valid surveys.

Cleaning Procedures for A3 – A19

Table 4-6 summarizes the allocation of missing responses for nine of the major questions in the Home and Lifestyle section before and after the cleaning process. The table shows that some variables that were missing prior to cleaning were assigned values during the process, while others were set to "not applicable." In addition, some survey responses were found to be logically inconsistent with other responses or utility-provided information and changed to missing, not applicable, or reallocated to a new response. The procedures used to cross-reference and clean these variables are discussed below.

Table 4-6: Cleaning Results for Missing Home and Lifestyle Responses

	Missing Pre-Cleaning			Assigned Value Pre-Cleaning			
Home & Lifestyle (A2 - A19)	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Unchanged Post Cleaning
Own or rent home (OWNRENT)	330	0	0	0	0	0	25,391
How long at address (YRS_RES)	233	0	0	0	0	0	25,488
Seasonal Occupancy (SEASOCC)	0	0	382	120	0	649	24,570
Year home built (BUILTYR)	1,726	0	0	371	0	0	23,624
Number of bedrooms (NUMROOM)	154	0	0	1,054	0	24,513	0
How many square feet of living space (SQFT)	2,096	0	0	329	0	0	23,296
Attic/Ceiling Insulation (ACEILINS)	3,655	0	790	0	0	30	21,246
Ceiling Insulation (CEILINCH)	3,236	9,188	0	0	0	0	13,297
Remodeling (REMOD)	640	0	238	0	0	35	24,808
Is natural gas available (NGSERV)	415	0	1,117	0	0	907	23,282
Natural gas hookup in home (NGLINE)	300	1,643	1,842	2	1,329	1,190	19,415

Source: *2010 California Residential Appliance Saturation Survey*

A3 – Years Respondent Lived in Home (YRS_RES)

Responses to questions A3 and A6 indicated how long a respondent has lived in the residence (*YRS_RES*) and the year that dwelling was built (*BUILTYR*). If the response to YRS_RES was greater than the age of the dwelling as indicated by BUILTYR, both variables were set to missing. In addition, the *BUILTYR* variable was cross-referenced with the age of the primary heating and water systems, *HTSAGE* and *PRWHAGE* respectively. If the age of the dwelling was less than the age of these systems, then *BUILTYR* and the respective system ages were set to missing.

A4/A5 – Seasonal Occupancy (SEASOCC) and (SEASJAN – SEASDEC)

While the 2003 RASS attempted to address partial year residents, the CDA in this study was estimated using only dwellings occupied year-round. (See Chapter 5 for a detailed discussion of the CDA process.) Respondents were asked whether the dwelling was occupied on a year-round or seasonal basis, or serves as a vacation residence (A4 – *SEASOCC*). Cleaning this variable was a two-step process.

First, responses for *SEASOCC* were checked against electric billing records to determine whether there was electric consumption for each month of the year. If a respondent indicated that the dwelling was a year-round residence, but the billing records showed zero electric consumption for three consecutive billing periods, then the dwelling was flagged as not occupied for part of the year. Conversely, households that showed electric consumption for all

SB_GT&S_0771223

billing periods, but the respondent indicated the dwellings was not a year-round residence were identified with a flag indicating a year-round resident.

The second step of this process checked responses for *SEASOCC* against survey variables reporting the months a respondent indicated living in the residence (A5-*SEASJAN* – *SEASDEC*). The following rules were used for this cross-reference:

- SEASOCC was set to 4 (vacation or rental home) for dwellings occupied for two or less months.
- SEASOCC was set to 2 (partial-year or seasonal residence) for dwellings occupied for three or more months.
- SEASOCC was set to 1 (year-round residence all months were left blank.

A6 – Year Dwelling Was Built (BUILTYEAR)

Responses for A6 (*BUILTYR*) were cleaned using the same logic as A3 (*YRS_RES)*. The variable *BUILTYR* was used to construct the variables *AGEHOME*, a continuous variable for the age of the dwelling, and *NEWHOME*, an indicator variable for new construction. The imputation of these variables is discussed further in the CDA data imputation section later in this chapter.

A7/A8 – Number of Bedrooms (NUMROOM) and Square Feet of Living Space (SQFT)

For a given dwelling type (DWLTYPE), the number of bedrooms (NUMROOM) was assumed to be constrained by the square footage (SQFT). The rules used to determine the logical consistency of these three variables are presented below. Cases in which the rules were violated, both the square footage and the number of bedrooms were set to missing.

- Single-Family Dwellings (DWLTYPE =1) –Less than 2,000 square feet with more than eight rooms; or less than 250 square feet.
- Townhouses (DWLTYPE = 2) – Less than 2,000 square feet with more than eight rooms; or less than 250 square feet with more than one room.
- Apartments (DWLTYPE = 3,4) –Less than 1,500 square feet with 4 or more rooms; or less than 250 square feet with more than one room.
- Mobile Homes (DWLTYPE = 5) – Less than 1,500 square feet with 4 or more rooms; more than five rooms; or less than 250 square feet with more than one room.

In addition, the *SQFT* variable was used to derive the continuous variable (*SQFT_A*). The *SQFT_A* variable is a continuous variable derived from the *SQFT*. These variables were used in the conditional demand analysis model, which required missing values to be imputed. The imputation of these variables is discussed further the CDA data imputation section later in this chapter.

A9 – Exterior Walls (EXTWLINS)

Responses to *EXTWLINS* were unchanged.

A10/A11 – Attic/Ceiling Insulation (ACEILINS and CEILINCH)

The variable *ACEILINS*, whether the attic or ceiling is insulated, was cross-referenced with *CEILINCH*, number of inches of insulation in the attic or ceiling. If the response to *ACEILINS* was either "no" or missing, but they provided the number of inches (i.e. *CEILIINCH* was not missing), then the response for *ACEILINS* was changed to "yes."

A13 – Remodeling (REMOD) and A14 – Type of Remodeling

Respondents were asked whether the home has been remodeled in the past 12 months (A13 – *REMOD*) and then asked to indicate the type of remodeling (A14). If a respondent skipped or answered "No" to A13, but indicated that a type of remodeling in A14, the response to A13 was changed to "yes."

A15 – Number of Occupants by Age Group (NR0-5, NR6-18, NR19_34, NR35-54, NR55-64, NR65-99)

The survey requested respondents to identify the number of individuals residing in the household according to six different age groups. Although response categories included a "ZERO" option, it is common for respondents to simply skip age groups that do not apply to their household. Therefore, the following criteria were used to distinguish between skipped responses that are not relevant and those that did not respond to the set of questions:

- If a respondent skipped all questions pertaining to the number of residents by age group, then all values were set to 97 or missing.
- If at least one category was filled out, then the age groups that did not have a response were set to zero.
- The total number of residents was also set to missing if all age groups were missing or zero.

The following variables were created during the cleaning process to be used in the CDA and cross-tabulations of survey responses:

- Number of people living in the household (RESCNT)
- Number of people living in the household over 65 (SENIORS)
- Number of people living in the household under 19 (KIDS)
- Number of people living in the household 19-64 (ADULTS)

Missing values of the *RESCNT* variable were imputed for the CDA analysis, creating the new variable *NUMI*, which will be discussed in the CDA data imputation section later in this chapter.

A18 – Natural Gas Availability

Responses to natural gas service being available in their area (*NGSERV*) that were either missing or reported as "no" were changed to "yes" if the cleaning process used for question A19 (*NGLINE*) indicated that they had a natural gas line to the residence.

A19 – Natural Gas Hookup in the Home (NGLINE)

Where possible, the presence of a natural gas line of the residence (A19 *NGLINE*) was checked against billing information sent by the three gas utilities based on the following:

- If the respondent indicated he or she did not have a gas line to the residence but was found to have gas billing records, then the response was changed to yes.
- If the respondent indicated having a gas line with service provided by one of the three major utilities, but no billing records were found, then the response was changed to no.
- If the respondent reported gas service provided by one of the smaller utilities, responses to the number of natural and bottled gas appliances were used to confirm the presence of a natural gas line to the residence.
- The new variable NGLINE2 recorded the corrected response to NGLINE after verifying the survey response.

Additional Cleaning of Your Home and Lifestyle Variables

A1 (subset) – Number of Stories (STORIES)

Respondents who live in single family dwellings (A1-*DWLTYPE*= 1) were instructed to answer this question, while those living in all other dwelling types were instructed to skip it. If a respondent from one of the other dwelling types provided an answer to *STORIES*, the response was changed to 99 (not Applicable).

A2 – Own or Rent Dwelling (OWNRENT)

Responses to the *OWNRENT* question are unchanged.

A20 – Number of Vehicles and A21 – Vehicle Descriptions

Question A20 (*NUMVEH*) asked respondents indicate the number of vehicles in the household, while Question 21 contained a set of variables that recorded characteristics of up to three of those vehicles. Vehicle characteristics included the following:

- Number of miles driven on one way regular trips per weekday
- Total miles driven per year
- Location that vehicles are parked
- Presence of an electric power outlet near the vehicle's parking area

The number of vehicles listed and vehicle characteristics were cross-referenced to ensure they were logically consistent. If they were found to be inconsistent, the cleaning process attempted to infer the correct response given the available information. For example, if a respondent filled in the vehicle characteristics, but the number of vehicles (*NUMVEH*) was missing or less than the number of vehicles for which they provide characteristics, then the number of vehicles was set to be consistent with the characteristics data. If *NUMVEH* was left blank and no characteristics were provided, then *NUMVEH* was set to no response (97).

In addition to checking the total number of vehicles, the characteristics of each vehicle (one, two, and three) were checked against each other. For each set of vehicle characteristics, if a respondent to skipped the information for the lower number vehicle (i.e. Vehicle 1) and populated the data for a higher vehicle number (i.e. Vehicle 2), then the characteristics were assumed to apply to the lower number vehicle. If the number of vehicles, *NUMVEH*, was larger than the set of vehicle characteristics provided, the characteristics were set to "missing."

Cleaning Space Heating Survey Responses

This section covers the procedures used to eliminate survey multiple responses and inconsistencies in responses in Section B – Space Heating of the survey. The cleaning process also revealed substantial fuel misreporting. Fuel misreporting is reviewed in the CDA variables section later in this chapter, which also discusses additional primary space heating system variables that were derived for the CDA model and data imputation.

B1 – Pay for Heat (PAYHEAT)

The question concerning how a household pays for heat (*PAYHEAT*) was critical to the process used to clean the remaining heating questions. The following cross-references were used to evaluate logical inconsistencies and make corrections wherever possible:

Only households indicating that they pay for heat directly were asked to fill out the majority of the heating questions, while those who either indicated heat was included in the rent or that they do not have a heating system were asked to only answer questions concerning portable electric heaters. If information was provided for at least one heating system and *PAYHEAT* was either *no* or *no – included in their rent* then a new variable *PAYHEAT1* recorded the response as *yes*. The original *PAYHEAT* variable was preserved by the original pay for heat response.

For cases in which multiple responses were provided, the lowest numbered response was kept, This logic favored “yes –pay for heat” over “no – it is part of my rent/condo fee.” Similarly, “no – included in rent” was chosen over “no – do not have heating system.”

If a respondent indicated that they do pay for heat, but did not list any heating systems, or *PAYHEAT* was missing and they did not list any heating systems then system variables were set to missing (97). If a respondent did not pay for heat then all heating system variables were set to not applicable (99).

Table 4-7: Heating Payment Question Cleaning

	Payheat	
	Total	% Total
Yes - Pay for Heat	22,228	90.9%
No - Included in Rent/Fee	559	2.3%
No - Do not Have Heating	928	3.8%
Missing	749	3.1%
All	24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

B2 – Type of Primary and Secondary Heating Systems

The type of primary and secondary heating system served as the basis for system fuel types and heating UEC estimates. This section deals with the rules used to confirm the consistency of heating systems with billing data and other survey variables. Specific rules used to clean natural gas, electric, propane, and other heaters are listed below.

Natural Gas Heating Systems

If a respondent indicated having a primary or secondary natural gas heating system, the cleaned variable a natural gas line to the home (NGLINE) was cross referenced. As discussed above, this confirmed the heating system was consistent with the survey response for NGLINE as well as billing information. If a dwelling did not have gas service, but indicated a natural gas system, the system response was set to missing (97). In addition, if a respondent reported having natural gas radiators, but either the radiators were not the primary heating system, or there was also a forced hot air system, then the radiators were not included as a heating system.

Electric Heating Systems

Survey responses for electric heating systems were checked to determine whether both a central heat pump heating and central forced air heating were indicated. If a respondent indicated having both types of systems, then the heat pump was selected as the primary heating system, and the central forced air system was set to missing. The survey allowed for up to six primary and six secondary electric heating systems. If a respondent reported having five or more electric space heaters, then all were set to missing.

Propane Heating Systems

If a respondent had natural gas in the residence and indicated having propane heat, the propane systems were set to missing.

Other Heating Systems

The number of "other" space heating systems was restricted to two systems. If respondents indicated having more than two "other" systems, then all systems were set to missing.

Primary and Auxiliary Heating Fuels

If a respondent provided more than one primary heating system, then the first system selected was set as the primary system and the subsequent responses were assigned to auxiliary heat. If only additional heating systems were provided, then the primary heat was set to the first additional heating system indicated.

The new variables primary heating fuel (*PHTFUEL*) and auxiliary heating fuel (*AHTFUEL*) were derived from the primary and additional heating system information. If the respondent indicated they had portable electric heaters, and did not have natural gas auxiliary heat, then *AHTFUEL* was set to 2 for electric heat. The coding used for *PHTFUEL* and *AHTFUEL* is summarized in Table 4-8 below.

Table 4-8: Primary Heating Fuel Data Cleaning

PHTFUEL	Code	Total	% Total
Natural Gas	**1**	17,380	71.0%
Electric	**2**	3,980	16.3%
Bottled Gas	**3**	847	3.5%
Wood	**4**	561	2.3%
Solar	**5**	0	0.0%
Other	**6**	43	0.2%
Missing (Respondent failed to answer question)	**97**	731	3.0%
Not Applicable (Respondent does not pay for heat or does not have a heating system)	**99**	922	3.8%
All		24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

B3 – Pilot Light for Primary and Secondary Natural Gas Heating System (MAINPLT and SECPILT)

The survey requested information on the use of pilot lights for primary and secondary natural gas heating systems, *MAINPILT* and *SECPILT*, respectively. If a respondent did not report having a natural gas system, then the *MAINPILT* and *SECPILT* were set to not applicable (i.e. 99). For respondents who had either a primary or secondary natural gas system and no response was given for *MAINPILT* or *SECPIL*, respectively, then the value was set to missing (97)

B6 – Heat Temperature Setting

Responses to the heating temperature settings were cross checked with information concerning how the household pays for heat, whether they have a heating system, and whether they skipped all temperature settings. Rules used to clean this section include:

- If respondents skipped PAYHEAT or indicated paying for heat but provided no indication regarding the type of heating system, then all temperature settings were set to missing (97).
- All temperature settings were set to Not Applicable (99) if they did not pay for heat, or if it is included in their rent, or if they do not have a thermostat.
- For respondents with a thermostat, temperature settings were evaluated to determine whether respondents answered a setting for at least one time of day. If they provided at least one setting, then all missing temperatures were set to off. If they did not provide any settings, then all were set to missing (97).

Space Cooling

This section covers the procedure used to eliminate survey multiple responses and inconsistencies in responses to survey Section B – Space Cooling. The space cooling section recorded information concerning central air conditioning and room air conditioning. Respondents were first asked how they pay for central air conditioning. Those who either did not pay for cooling or indicated it was included in their rent, were asked to skip to the room air conditioning section.

C1 – Pay for Cooling

How a household pays for cooling (*PAYCOOL*) was evaluated similarly to the *PAYHEAT* variable, and is summarized in Table 4-9. For households that did not indicate they pay for cooling, but provided information on cooling systems, a new variable *PAYCOOL1* recorded the response as "yes." For cases in which multiple responses were provided, the lowest numbered response was kept. If respondents indicated that they do pay for cooling, but did not list any central cooling systems, or *PAYCOOL* was missing and they did not list any cooling systems then system variables were set to missing (97). If a respondent did not pay for cooling, then all cooling system variables were set to not applicable (99).

Table 4-9: Cooling Payment Question Cleaning

PAYCOOL	Code	Total	% Total
Yes - Pay for Cooling	1	13,213	54.0%
No - Included in Rent/Fee	2	481	2.0%
No - Do not Have Cooling	3	9,691	39.6%
Missing	97	1,079	4.4%
All	All	24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

C2 – Central AC

The following checks were used to clean the central air conditioning section:

- Respondents that reported the central air conditioner, evaporative cooler, or heat pump was zoned, but did not indicate the number of the respective systems, were assigned one system of that type.

- If they have a central heat pump for heating, and they indicated having central air conditioning, then they were assigned a central heat pump.
- If the survey indicated the addition of a central air conditioning unit in the past 12 months and the household owns the dwelling, yet there are no central air conditioning units specified, then the number of central air conditioning units was set to one.

C3 – Age of Central Air Conditioner (CLCNTAGE)

The variable for age of central air conditioner was cross referenced with the presence of a cooling system. If a system was reported, but no age was provided then age was set to missing.

C4 –Central Air Conditioner Temperature Setting

Responses to the central air conditioner temperature settings were cross referenced with information concerning how the household pays for cooling, whether they have a central air conditioning system, and whether they skipped all temperature settings. Rules used to clean this section include:

- If respondents skipped PAYCOOL or pay for heat but provided no indication of the type of cooling system, then all temperature settings were set to missing (97).
- All temperature settings were set to Not Applicable (99) if they did not pay for central cooling, or central cooling is included in their rent, or if they do not have a thermostat.
- For respondents with a thermostat, temperature settings were evaluated to determine whether respondents answered a setting for at least one time of day. If they provided at least one setting, then all missing temperatures were set to off. If they did not provide any settings, then all were set to missing (97).

C7 – Room AC

The first step in cleaning the room air condition questions was to ensure that information for the first air conditioner was populated before information for the second room air conditioner. If a respondent did not populate the first air conditioners information, but provided responses for the second, this information was moved to the first unit. Similarly, if the second air conditioner's information was left missing, but the third was populated, then this datum was moved to the second. This was done for both the type and age of each air conditioner.

Next, the total number of room air conditioners was determined and used for the following checks:

- If at least one unit was present in the residence, then the variable NOROOMAC was set to zero. If no units were found then NOROOMAC was set to one.
- If only one unit was found, then all information for the second and third units were set to not applicable (99).

- If the survey indicated the addition of a room air conditioning unit in the past 12 months (WWADD=1) but no room air conditioning units were specified, one was added and NOROOMAC was set to 0.

The type of room air conditioner was cross referenced with the age of each unit. If the respondent filled in an age but left the type blank, then the unit was assumed to be a window/wall air conditioner. However, if type of room air conditioner was indicated, but age was left blank, then age was set to missing (97).

Cleaning Water Heating Survey Responses

This section covers the procedure used to eliminate multiple and inconsistent responses to survey Section D – Water Heating. The cleaning process also revealed substantial fuel misreporting as reviewed in the CDA data imputation section later in this chapter.

D1 – Pay for Water Heat (PAYWH)

The question concerning how a household pays for heat (*PAYWH*) serves as the basis for cleaning the remainder of the section because households that indicated they do not pay for water heat, directly, were asked to skip this section. If information was provided for at least one water heater and *PAYWH* was either "no" or "no – included in their rent" water then a new variable *PAYWH1* recorded the response as "yes." The original *PAYWH* variable was preserved by the original pay for water heat response.

For cases in which multiple responses were provided, the lowest numbered response was kept. This logic favored "yes –pay for water heat" over "no – it is part of my rent/condo fee." Similarly, "no – included in rent" was chosen over "no – do not have water heating system."

If respondents indicated that they do pay for water heat, but did not list any heating systems, or *PAYWH* was missing and they did not list any water heating systems then system variables were set to missing (97). If a respondent did not pay for water heat, then all water heating system variables were set to missing (99).

Table 4-10: Water Heating Payment Question Cleaning

PAYWH	Code	Total	% Total
Yes - Pay for Water Heating	1	20,497	83.8%
No - Included in Rent/Fee	2	2,833	11.6%
No - Do not Have Water Heating	3	349	1.4%
Missing	97	785	3.2%
All	All	24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

D2 – Type of Primary and Secondary Water Heating Systems

The type of primary and secondary water heating system served as the basis for system fuel types and water heating UEC estimates. This section deals with the rules used to confirm the consistency of water heating systems with billing data and other survey variables. Specific rules used to clean natural gas, electric, propane, and other heaters are listed below.

Natural Gas Water Heating Systems

For respondents who indicated they have a primary or secondary natural gas water heater, the cleaned variable a natural gas line to the home (*NGLINE*) was cross-referenced. As discussed above, this confirmed that the water heater was consistent with the survey response for *NGLINE* as well as billing information. If a household did not have gas service, but the survey response indicated a natural gas water heater, the system response was set to missing (97).

Propane Water Heaters

Households that had natural gas (*NHLINE*=1) were not allowed to have a propane water heater. For these households, all propane water-heating systems were set to missing.

Solar Water Heaters

Solar water heaters are only allowed for single family dwellings. Responses for all other dwelling types that indicated solar were set to missing (97).

Other Water Heaters

Cases in which more than two "other" water heaters were indicated were set to missing.

Primary and Auxiliary Water Heating Fuels

If a respondent provided more than one primary heating system, then the first system selected was set as the primary system, and the subsequent responses were assigned to auxiliary heat. If only additional heating systems were provided, then the primary heat was set to the first additional heating system indicated. The maximum number of water heaters was set to four, such that surveys with five or more types of water heaters were considered erroneous. For these surveys, responses to all water heater type questions were set to missing (97).

The new variables primary heating fuel (*PRWHFUEL*) and auxiliary heating fuel (*AWHTFUEL*) were derived from the primary and additional heating system information. If the respondent indicated having portable electric heaters and did not have natural gas auxiliary heat, then *AWHTFUEL* was set to 2 for electric heat. Coding for *PRWHFUEL* is shown in Table 4.11

Table 4-11: Water Heating Fuel Data Cleaning

PRWHFUEL	Code	Total	% Total
Natural Gas	1	17,618	72.0%
Electric	2	2,375	9.7%
Bottled Gas	3	1,135	4.6%
Solar	4	9	0.0%
Other	5	18	0.1%
Missing	97	982	4.0%
N/A	99	2,327	9.5%
All		24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

D5 – Number of Showers and Baths Per Day

A limit was set on the number of showers/baths taken per day (*SHWRDAY or BATHDAY*) based on the cleaned number of residents (*RESCNT*). This limit was set at two showers or baths per day per person. Responses for households that exceeded this limit were set to missing (97).

Laundry

This section covers the procedures used to clean Section E – Laundry Equipment. The cleaning process involved correcting for fuel misreporting on clothes dryers, as reviewed in the CDA data imputation section later in this chapter.

E1 Presence of Laundry Equipment in Home (LNDRYEQU)

Question E1 (*LNDRYEQU*) asked whether laundry equipment was present in the home, not present, or was in a common area. Only respondents with laundry equipment in the home were instructed to fill out the remainder of the section. Therefore, responses were cross referenced with responses to (E2) – Clothes Washer Type (*CWTYP*), (E3) – Clothes Washer Age (*CWAGE*), and (E5) clothes dryer type (*CDTYP*). If a respondent answered any of the questions pertaining to the type of laundry equipment in the home and answered, then *LNDRYEQU* was changed to "Yes."

E4 – Number of Clothes Washer Loads per Week

A limit was set on the number of loads washed per average week (sum of *CWHWLD, CWWWLD, CWCWLD*) based on the cleaned number of residents (*RESCNT*). This limit was set at five loads per week per person. Responses from households that exceeded this limit were set to missing (97).

E5 – Clothes Dryer Type

For respondents who indicated they have a primary natural gas dryer, the cleaned variable, a natural gas line to the home (*NGLINE*) was cross referenced. As discussed above, this confirmed the clothes dryer was consistent with the survey response for *NGLINE* as well as billing information. If a household did not have gas service, but the survey response indicated a natural gas dryer, the system response was set to missing (97).

E6 – Number of Clothes Dryer Loads per Week

A limit was set on the number of loads dried per average week based on the cleaned number of residents (*RESCNT*). This limit was set at five loads per week per person. Responses from households that exceeded this limit were set to missing (97).

Food Preparation

For respondents who indicated they have a primary natural gas range or oven, the cleaned variable, a natural gas line to the home (*NGLINE*), was cross-referenced. As discussed above, this confirmed that the cooking equipment was consistent with the survey response for *NGLINE* as well as billing information. If a household did not have gas service, but the survey response indicated a natural gas range or oven, the system response was set to missing (97).

Refrigerators

Question G1 (*RFNUM*) asked respondents to indicate the number of refrigerators they own, while Question G2 contained a set of variables that recorded characteristics of up to three of those refrigerators. Refrigerators characteristics included the following:

- Door style (RF1STY, RF2STY, RF3STY)
- Cubic feet (RF1SZ, RF2SZ, RF3SZ)
- Frost free or manual defrost (RF1DEF, RF2DEF, RF3DEF)
- Age (RF1AGE, RF2AGE, RF3AGE)
- Other features (RF1OTH, RF2OTH, RF3OTH)

The number of refrigerators listed and refrigerator characteristics were cross-referenced to ensure they were logically consistent. If they were found to be inconsistent, the cleaning process attempted to impute the correct response given the available information. For example, if a respondent filled in the refrigerator characteristics, but the number of refrigerators (*RFNUM*) was missing or less than the number of refrigerators for which they provide characteristics, then the number of refrigerators was set to be consistent with the characteristics data. If *NUMREF* was missing and no characteristics were provided, then *RFNUM* was set to missing.

In addition to checking the total number of refrigerators, the characteristics of each refrigerator (one, two, and three) were checked against each other. For each set of refrigerator characteristics, if a respondent skipped the information for the lower number refrigerator (i.e. Refrigerator 1) and populated the data for a higher refrigerator number (i.e. Refrigerator 2), then the characteristics were assumed to apply to the lower number refrigerator. If the number of refrigerators, *RFNUM*, was larger than the set of refrigerator characteristics provided, the characteristics were set to "missing."

Table 4-12 summarizes the allocation of missing responses for the refrigeration section.

Table 4-12: Missing Refrigerator Number and Characteristics

		Missing Pre-Cleaning			Assigned Value Pre-Cleaning			
Refrigerator number	Refrigerators (G1 - G2)	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Unchanged Post Cleaning
N/A	How many refrigerators do you have plugged in? (RFNUM)	47	0	274	0	0	25,270	130
RF1	Door Style (RF1STY)	444	33	161	0	0	0	25,083
	Size in Cubic Feet (RF1SZ)	1,645	33	187	0	0	0	23,856
	Frost-Free or Manual Defrost (RF1DEF)	1,763	33	135	0	0	0	23,790
	Age (RF1AGE)	799	33	135	0	0	0	24,754
	Other Features (RF1OTH)	299	33	17,045	0	0	0	8,344
RF2	Door Style (RF2STY)	509	19,419	11	121	59	4	5,598
	Size in Cubic Feet (RF2SZ)	619	19,451	11	145	27	3	5,465
	Frost-Free or Manual Defrost (RF2DEF)	856	19,457	12	99	21	1	5,275
	Age (RF2AGE)	416	19,443	17	86	35	8	5,716
	Other Features (RF2ITH)	590	19,474	4,989	26	4	302	336
RF3	Door Style (RF3STY)	168	25,035	0	19	8	0	491
	Size in Cubic Feet (RF3SZ)	155	25,034	0	19	9	0	504
	Frost-Free or Manual Defrost (RF3DEF)	207	25,034	0	20	9	0	451
	Age (RF3AGE)	138	25,014	0	12	29	0	528
	Other Features (RF3OTH)	185	25,042	468	1	1	18	6

Source: *2010 California Residential Appliance Saturation Survey*

Because the refrigeration variables were used to construct engineering estimates of monthly kWh consumption used in the CDA, missing values were also imputed. The data imputation and the engineering estimates are described in the CDA section later in this chapter.

Freezers

Question F1 (*FZNUM*) asked respondents to indicate the number of freezers they own, while Question F2 contained a set of variables that recorded characteristics of up to two of those freezers. Freezers characteristics included the following:

- Door style – (FZ1STY, FZ2STY)
- Cubic feet – (FZ1SZ, FZ2SZ)
- Age – (FZ1AGE, FZ2AGE)

The number of freezers listed and freezer characteristics were cross referenced to ensure they were logically consistent. If they were found to be inconsistent, the cleaning process attempted to impute the correct response given the available information. For example, if a respondent filled in the freezer characteristics, but the number of freezers (*FZNUM*) was missing or less than the number of freezers for which they provided characteristics, then the number of freezers was set to be consistent with the characteristics data. If *FZNUM* was missing and no characteristics were provided, then *FZNUM* was set to missing.

In addition to checking the total number of freezers, the characteristics of each freezer were checked against each other. For each set of freezer characteristics, if a respondent skipped the information for the lower number freezer (i.e. Freezer 1) and populated the data for a higher freezer number (i.e. Freezer 2), then the characteristics were assumed to apply to the lower number freezer. If the number of freezers, *FZNUM*, was larger than the set of freezer characteristics provided, the characteristics were set to "missing."

Table 4-13 summarizes the allocation of missing responses for the freezer section.

Table 4-13: Missing Freezer Number and Characteristics

Freezer number	Freezers (H1 - H2)	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Unchanged Post Cleaning
N/A	How manyfreezers do you have plugged in? (FZNUM)	1,204	0	362	0	0	23,986	169
FZ1	Door Style (FZ1STY)	1,410	19,522	24	0	0	0	4,765
	Size in Cubic Feet (FZ1SZ)	1,641	19,522	28	0	0	0	4,530
	Age (FZ1AGE)	1,515	19,522	23	0	0	0	4,661
FZ2	Door Style (FZ2STY)	1,246	24,258	0	12	12	0	193
	Size in Cubic Feet (FZ2SZ)	1,269	24,256	0	14	14	0	168
	Age (FZ2AGE)	1,250	24,255	0	8	15	0	193

Source: *2010 California Residential Appliance Saturation Survey*

Because the freezer variables were used to construct engineering estimates of monthly kWh consumption used in the CDA, missing values were also imputed. The data imputation and the engineering estimates are described in the CDA section later in this chapter.

Spas and Hot Tubs

Only respondents who indicated they pay for the use of a spa or hot tub were asked to complete the remainder of this section. If respondents indicated they have a spa or hot tub in a common area, or do not have a spa or hot tub, but filled in information provided by questions I2 – I7, the response to I1 (*SPATYP*) was changed to "Yes, I pay for its energy use."

For respondents who indicated they have a natural gas spa heater, the cleaned variable a natural gas line to the home (*NGLINE*) was cross-referenced. As discussed above, this confirmed whether the spa water heater was consistent with the survey response for *NGLINE* as well as billing information. If a household did not have gas service, but indicated a natural gas spa heater, the system response was set to missing (97)

Respondents who lived in apartments were restricted from having a spa or hot tub.

Swimming Pools

Only respondents who indicated they pay for the energy use of a swimming pool were asked to complete the remainder of this section. If a respondent indicated they have a pool in a common area, or do not have a pool, but filled in information provided by questions J2 – J7, the response to I1 (*PLTYP*) was changed to "Yes, I pay for its energy use."

For respondents who indicated they have a natural gas pool heater, the cleaned variable a natural gas line to the home (*NGLINE*) was cross-referenced. As discussed above, this confirmed whether the pool heater was consistent with the survey response for *NGLINE* as well as billing information. If a household did not have gas service, but indicated a natural gas pool heater, the system response was set to missing (97)

Respondents who did not live in single-family dwellings were restricted from having a pool.

Entertainment and Technology

Responses to the entertainment and technology section were evaluated to determine whether respondents skipped appliances they do not have or skipped all questions. In particular, if respondents answered at least one technology question, then all missing values were set to zero. If they did not provide a response to any technologies, then all were set to missing (97). Cleaning of these variables is summarized in Table 4-14.

Additional variables were also constructed and missing values were imputed for the CDA, as discussed in the CDA section later in this chapter.

Table 4-14: Missing Entertainment and Technology Appliances

	Missing Pre-Cleaning			Assigned Value Pre-Cleaning			
Entertainment and Technology Question (K1)	**Missing Post Cleaning**	**Not Applicable Post Cleaning**	**Re-Assigned Post Cleaning**	**Missing Post Cleaning**	**Not Applicable Post Cleaning**	**Re-Assigned Post Cleaning**	**Unchanged Post Cleaning**
Standard Television (STDTV)	197	0	5,249	0	0	20,275	0
LCD television, smaller than 36 inches (SMLLCDTV)	197	0	12,369	0	0	13,155	0
LCD television, 36 inches or larger (LRGLCDTV)	197	0	12,318	0	0	13,206	0
Plasma television (PLSMTV)	197	0	15,311	0	0	10,213	0
Converter box for standard TV (DTA)	197	0	14,163	0	0	11,361	0
Cable or satellite bow without DVR (BOXNODVR)	197	0	11,644	0	0	13,880	0
Cable of satellite box with DVR (BOXWDVR)	197	0	11,972	0	0	13,552	0
Stand-alone digital video recorders (SADVR)	197	0	15,084	0	0	10,440	0
DVD player and/or VCR (DVDVCR)	197	0	5,953	0	0	19,571	0
Gaming systems (GAMSYS)	197	0	12,688	0	0	12,836	0
Separate sound or stereo system connected to TV (AUDTOTV)	197	0	11,871	0	0	13,653	0
Stand-alone stereo, I-pod of MP3 docking station (SAMUSIC)	197	0	12,077	0	0	13,447	0

Source: *2010 California Residential Appliance Saturation Survey*

Lighting

The lighting section consisted of a set of questions to gather information on both interior and exterior lighting.

Cleaning of the interior lighting section differs from the 2003 RASS because the section changed significantly since the 2003 study. Question L1 recorded the number of compact florescent light bulbs (CFLs) and incandescent light bulbs by room. If a respondent did not provide an answer to the number of CFLs or incandescent light bulbs for a given room, the variable was coded as missing (97). Since the survey responses were coded such that a value of one was equal to zero bulbs, and two was equal to one bulb, all non-missing responses were given a response value equal to one minus the survey response number. For example, if they provided a response value equal to one, they were coded as one minus one, or zero. This is because the first response was *zero.* This logic also was used for incandescent lights.

SB_GT&S_0771238

Question L2 recorded the number of interior lights used by time of day. Responses to L2 were cleaned using the same logic as L1. If a respondent did not provide an answer for a given time period, the variable was coded as missing (97).

Question L3 asked about interior lighting products, such as timers, sensors, and dimmers. If a respondent did not provide an answer for a given product, the variable was coded as missing (97). All valid responses were coded as the response value minus one.

The cleaning procedures used for exterior lights (L4) varied from those used for interior lighting. The process used to clean this section was consistent with the cleaning process used in the 2003 RASS because a specific UEC was estimated for exterior lighting. Responses to all exterior lighting products were examined to identify missing values. If the all values were skipped, then each value was coded as missing (97). If at least one value was provided for one of the products, then missing values were set to zero. The CDA model required additional lighting variables for exterior lighting, as discussed in the CDA section later in this chapter.

In addition, if a respondent skipped the question CFLTOINCD (L5), it was coded as 97; otherwise it was left unchanged.

Miscellaneous Appliances

Responses to the M1, number of miscellaneous appliance used, were evaluated to determine whether respondents skipped appliances they do not have or skipped all questions; these are summarized in Table 4-15. In particular, if a respondent answered at least one appliance questions, then all missing values were set to zero. If they did not provide a response to any appliance questions, then all were set to missing (97).

Additional variables were also constructed and missing values were imputed for the CDA, as discussed in the CDA data imputation section later in this chapter. Table 4-6 presents the percent of responses with missing values for the variables used to develop the appliance ownership indicator variables for the CDA.

Table 4-15: Missing Miscellaneous Appliances

	Missing Pre-Cleaning			Assigned Value Pre-Cleaning			
Miscellaneous Appliances (M1)	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Unchanged Post Cleaning
Chargers left plugged in all the time (CHRGRS)	693	0	1,739	0	0	23,289	0
Portable Fan (FNPORT)	693	0	2,239	0	0	22,789	0
Ceiling Fan (FNCEIL)	693	0	1,676	0	0	23,352	0
Attic Ventilator (WNDATV)	693	0	3,552	0	0	21,476	0
Electric Attic Fan (FNATTIC)	693	0	3,350	0	0	21,678	0
Whole House Fan (FNWHOLE)	693	0	3,429	0	0	21,599	0
Electric Air Cleaner (AIRCLEAN)	693	0	3,346	0	0	21,682	0
Humidifier or Dehumidifier (HUMDEH)	693	0	3,376	0	0	21,652	0
Wine or beverage cooler (WINCLR)	693	0	3,321	0	0	21,707	0
Water Purification (WHPURIFY)	693	0	3,189	0	0	21,839	0
Domestic hot water recirculation pump (DHWRPMP)	693	0	3,462	0	0	21,566	0
Electric Blanket (ELBLNKET)	693	0	2,984	0	0	22,044	0
Aquarium (AQUAR)	693	0	3,295	0	0	21,733	0
Trash Compactor (TRSHCOMP)	693	0	3,236	0	0	21,792	0
Sauna – Electric (SAUNA)	693	0	3,446	0	0	21,582	0
Electronic Security System (SCRTYSYS)	693	0	3,173	0	0	21,855	0
Pool or water garden Pump (POND)	693	0	3,287	0	0	21,741	0
Electric Garage Door Opener (GRGDROPN)	693	0	2,109	0	0	22,919	0
Lawn Mower – electric (LAWNMOWR)	693	0	3,362	0	0	21,666	0

Source: *2010 California Residential Appliance Saturation Survey*

Household Information

The variables *PTHME* (N1), *PTHMELOC,* and *PTHMEUTL* (N2) were cross-referenced for cleaning. If the respondent left vacation home (*PTHME*) blank, but filled in data for location or electricity provider, its value was changed to *yes.*

The household income variable was used to create the variable *AVGINC,* which was used in the CDA analysis. A summary is provided in Table 4-16.

Table 4-16: Missing Household Information

	Missing Pre-Cleaning			Assigned Value Pre-Cleaning			
Household Information (N1 - N7)	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Missing Post Cleaning	Not Applicable Post Cleaning	Re-Assigned Post Cleaning	Unchanged Post Cleaning
Own vacation home (PTHME)	1,204	0	27	0	0	184	24,306
Location for vacation home (PTHMELOC)	1,397	22,609	0	0	50	0	1,665
Electric utility provide for vacation home (PTHMEUTL)	1,377	22,401	0	0	258	0	1,685
Highest level of education (EDUC)	1,099	0	0	0	0	0	24,622
Primary spoken language (ETHNIC)	914	0	0	0	0	0	24,807
Number of occupants of home disabled (DISABLED)	1,215	0	0	0	0	0	24,506
Household total annual income (INCOME)	3,218	0	0	0	0	0	22,503

Source: *2010 California Residential Appliance Saturation Survey*

Billing and Weather Data

This section discusses the development of the data that were stored in the RASS billing database. This section includes a description of the billing databases provided by each utility, the methods used to clean the billing data, the normalization routines employed to standardize the consumption amounts and the merging of the billing data with the survey data. The primary data used from the utility billing databases were consumption amounts, read dates, and read codes. Dollar amounts of bills and payment information was not used in the analysis or included in the datasets provided. For simplicity, the authors refer to the data and databases as "billing data."

Billing Databases

Billing data were requested from the three California IOUs (Pacific Gas and Electric Company, Southern California Edison and San Diego Gas & Electric Company) and the Los Angeles Department of Water and Power for all sampled households. The section below describes the contents of the billing data from the utilities. In addition, gas consumption data were also requested from Southern California Gas Company for all electric respondents that could be matched to a gas account using an account matching process.

Pacific Gas and Electric Company

The gas and electric billing data for PG&E respondents were provided in a SAS data set. The billing data for PG&E respondents contained information from 35,266 electric account IDs and 25,335 gas account IDs for 35,356 unique premises. PG&E's billing data included the following set of information: premise identification number, account identification number, service agreement ID, gas and electric rate schedules, prior and current read dates, kWh and therm consumptions, gas and electric tariffs, residential dwelling types, customer information, the number of days in along with a start and end date for the billing period. Both the electric and gas billing data covered the period April 2008 through September 2009. A separate file identifying PG&E's net-metered customers was also provided.

Southern California Edison

The billing data for SCE were provided in three text datasets. These text data sets contained a total of 61,561 unique service account ids within the SCE territory. SCE's billing data included the following set of information: a customer number, customer name, address, a premise number, meter number, kWh consumption, bill date, number of billing days, tariff, Energy Commission weather zone, and SCE weather stations. The billing data covered the period June 2008 through September 2009. A separate file identifying SCE net-metered customers was also provided.

San Diego Gas & Electric Company

SDG&E Energy Commission-Sample

The SDG&E billing data contained data for 8,730 premises within SDG&E's territory. The SDG&E gas and electric data were provided in a single data set. The data set had billing data for 8,730 households with electric consumption and 5,639 households with gas consumption. The data included a premise identification number, customer number, service point ID, customer name, service and mailing addresses, service type code, read date, electricity tariff, gas tariff, the number of days in along with a start and an end date for the billing period, electricity consumption, therm consumption, and the number of billing days in the cycle. The billing data covered the period May 2008 through September 2009. A separate file identifying the net-metered customers was also provided.

SDG&E Oversample

The SDG&E Oversample billing data contained data for additional 11,291 premises within the SDG&E territory. The oversample data were also provided in a single data set. This data set had billing data for 11,291 households with electric consumption and 7,645 households with gas consumption. The data included the same fields as the SDG&E data described above. In addition, it included a flag that identified net-metered customers. The billing data covered the period February 2006 through February 2010.

Los Angeles Department of Water and Power

The LADWP billing data were provided in two SAS datasets – a dataset on new customers and another dataset on all other customers. The data on new customers contained data for 1,898 accounts. The dataset on all other customers contained data on 13,615 accounts within the LADWP territory. The LADWP data included the account number, 14 billing read dates - readdate1,2,..,14, associated consumptions and the number of days in the billing cycle. The general LADWP billing data covered the period February 2005 through September 2009. A separate file identifying the net-metered customers was also provided.

Southern California Gas Company

The sample frame for the RASS study was developed from the residential electric population from the three IOUs and LADWP. As such, collecting natural gas billing data for respondents served by SoCal Gas involved a customer matching procedure between the RASS sample frame data and the SoCal Gas residential population. This procedure required each of the steps discussed below.

Step 1—Identify SoCal Gas ZIP Codes. The sample frame was sorted by ZIP code and merged with a file that contained the natural gas utility serving each ZIP code in California. The sample having SoCal Gas as the gas utility was saved for further analysis.

Step 2—Disaggregate Customer Address: The service address variable in the sample frame was disaggregated into the following pieces:

- Street number and number fraction
- Street direction
- Street name
- Apartment or unit number
- ZIP code

Code was developed for each utility that created the six pieces of the address. These pieces, along with the customer name and account information were matched against SoCal Gas' population data for further analysis.

Step 3—Customer Address Merging with SoCal Gas Master File: The merging of sample addresses with SoCal Gas master file data to capture account number and rate information involved several phases. The two files were first merged by ZIP code, street number, street number fraction, street direction, street name, and apartment/unit number to obtain the exact address matched cases in the first phase.

For the remaining unmatched sample, the second phase involved merging the files by ZIP code, street number, street number fraction, and street name followed by a case-by-case inspection to select matches. In Phase 2, master metered accounts were located along with addresses that may have a missing street direction or different apartment/unit number designation (e.g., D instead of 4). The customer name appearing in the sample frame as well as the SoCal Gas master file was utilized in this phase to select the appropriate record.

For the remaining unmatched sample after Phases 1 and 2, the third phase involved merging the files by ZIP code, street number, street number fraction, and the first six characters of the street name followed by a case-by-case inspection to select matches using the same approach as was described in Phase 2.

For the remaining unmatched sample after phases 1 through 3, the final phase involved merging the files by ZIP code and customer last name followed by a case-by-case inspection to select matches that may have slightly different street name spellings between data sources.

Step 4—Download SoCal Gas Account Number File: The merged records were placed into a SAS file and downloaded from the SoCal Gas mainframe. Finally, 1,821 PG&E, 13,850 LADWP, 43,495 SCE, and 718 SDG&E accounts were associated with service addresses that were matched to SoCal Gas accounts.

Step 5—SoCal Gas Billing Data: The SoCal Gas Account Number File was merged with the billing data provided by SoCal Gas. The billing data contained information on the addresses associated with the 54,230 RASS sample population. The billing data included a business account ID, customer name, customer address, read date, dwelling type, tariff rate code, and the therm consumption. SoCal Gas' billing database included natural gas consumption from March 2008 to September 2009.

All respondents in the study sample had electric billing account information, but not all had gas service. The coincidence for the two services is shown in Table 4-17.

Table 4-17: Comparison of Gas and Electric Utility Providers for Survey Respondents

Gas Provider	Electric Utility				All
	SDG&E	LADWP	PG&E	SCE	
PG&E	0	0	5,290	0	5,290
SoCal Gas	200	2,455	404	7,969	11,028
SDG&E	2,774	0	0	0	2,774
All	2,974	2,455	5,694	7,969	19,092

Source: *2010 California Residential Appliance Saturation Survey*

Table 4-18 shows that not all respondents who received electricity from PG&E and SDG&E received gas service from the same provider. SoCal Gas was found to serves 400 respondents of PG&E and 200 respondents of SDG&E. Among those households, 912 had individually metered electric service but master-metered gas service. The natural gas master metered households were not included in the gas CDA or gas degree-day normalization (DDN) modeling discussed below.

Table 4-18: Comparison of Gas and Electric Utility Providers for Survey Respondents with Individually Metered Gas and Electric Service

Gas Provider	Electric Utility				Total
	SDG&E	LADWP	PG&E	SCE	
PG&E	0	0	5,283	0	5,283
SoCal Gas	192	2,211	387	7,338	10,128
SDG&E	2,769	0	0	0	2,769
Total	2,961	2,211	5,670	7,338	18,180

Source: *2010 California Residential Appliance Saturation Survey*

Billing Data Cleaning and Preparation for Analysis (Weather-Normalization)

Billing data were received from five utilities (LADWP, PG&E, SDG&E, SCE, and SoCal Gas) for each household in the sample invited to take the survey. These data included at least one year worth of meter-read dates, energy billed, days in the billing period, tariff class or rate code, and whether the consumption was estimated or adjusted. Where available, information was requested on dwelling type (e.g., single-family shared-wall, single-family detached). As was discussed in the billing databases section, each electric utility also identified net-metered households (i.e. households that installed home generation of wind or solar energy).[1]

Prior to analysis, these data were combined and sorted, and anomalous or problematic billing series were flagged for possible exclusion from some of the analysis. Additional quality control

1 Net-metered customers offset purchased kWh consumption with power generated on site. This on-site generation would have an unpredictable impact on the CDA, so net metered customers were removed from the analysis.

tests were performed to ensure that each respondent was correctly and uniquely matched to a sequence of billing periods, or if they had both gas and electric service, to ensure that they were matched correctly to both. Cleaned and validated billing data were then prepared for weather normalization.

DDN provided a way to generate a household's consumption for a standardized year's weather. This normalization accomplished two things: First, it converted consumption series that span varying numbers of days to a one-year period. Second, it provided annual consumption for long-run normal weather conditions. This normalization facilitated comparisons across climate zones despite any unusual weather events that might have occurred in certain zones.

The normalization modeled monthly electric or gas consumption individually for each household. Each household-level electricity model was a linear function of heating degree-days and cooling degree-days, with respect to heating reference temperature estimated specific to the location of the household. Each household level gas model was a linear function of heating degree-days, with respect to heating reference temperature estimated specific to the household location.

Preparing the data for the analysis required the following tasks:

Separate Electric and Gas

For the two utilities that supply both electricity and natural gas (PG&E, SDG&E), the billing series for electric and gas services were separated. The weather-normalized annual consumption were calculated separately for the two fuels.

Identify Potentially Problematic Billing Periods

The next step was to identify any billing periods or series that might be problematic. For example, if a series included re-bills or balance bills, there would appear to be two records for the same service dates. The utilities provided various other flags to mark potential data issues such as multiple rate codes. Anomalies in billing data were examined including out of range read dates and extreme consumption quantities including zero or negative consumption billed. Flags were created to describe the resolutions, so that subsequent analysis and modeling could be checked for undue influence from these anomalies.

Where multiple meters matched a single respondent ID, the data were explored to determine whether they were serving a single household or multiple households. In the former case, monthly amounts were aggregated to reflect the household's total energy consumption, and in the latter, information from the survey responses was used to match the respondent to a single residence.

Data were then rearranged so that each record contained all billing periods for a single household in chronological order. Flags were created to mark short billing series (fewer than twelve billing periods), or short time period (less than one year). Some utilities identified recently established customers from longer-term ones, but the flags compared well with these customers. There were no billing series found to have billing periods missing in the middle of the series.

Prepare Data for Analysis

Data files from the different utilities were harmonized so the files could be combined into a single electric and a single gas billing file. This required standardizing variable names across the four electric and the three gas utility files. The two combined files were then sorted and transposed so that each record represented a single household, with all of the billing periods in chronological order. Each billing period is described by three key variables: read date, number of days of service in that period, and kWh or therms used during that period. Data received from the utilities included, at most, 21 electric billing periods and 20 gas billing periods for a single account. For each household, the first billing period was Period 1. If there were fewer than the maximum number of billing periods, higher-numbered periods were reported as missing. Each record in the combined billing data sets was then matched to its correct climate zone (T24 zone) and weather series, and to the corresponding survey data.

Weather-Normalization of Billing Data

To ensure that the analysis was comparable across utilities and climate zones, it was necessary to analyze the respondent energy consumption while controlling for the local weather, also called normalization. Techniques used for normalizing consumption include calendarization, selecting a specific time period for analysis that minimizes extreme weather, and modeling the relationship between weather and energy consumption. A modeling process was used to normalize consumption in the current study.

The normalization process used was the Degree-Day Normalization (DDN) similar to the *Princeton Scorekeeping Model* (PRISM™) technique. This method consists of two parts:

- Modeled each household's energy consumption as a function of outdoor temperature over the study period.
- Used each household's fitted model to calculate energy consumption for a year of standardized temperatures.

The results of the process provided Normalized Annual Consumption (NAC) estimates for each household. These NACs reflected the households' estimated energy consumption for a typical (normal) year.

Temperature Data

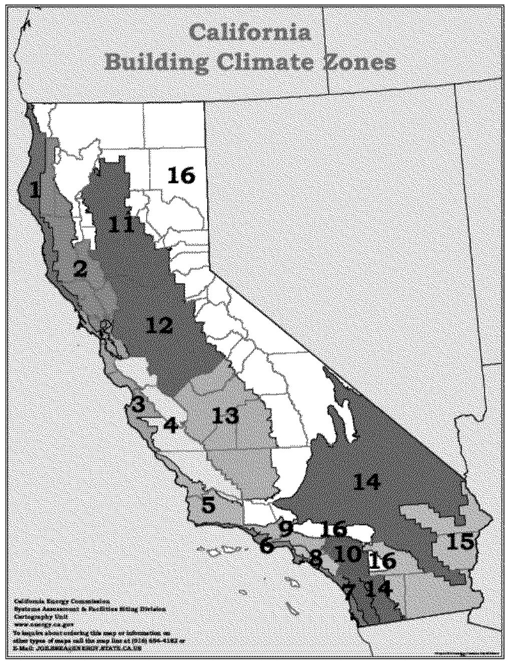
Separate temperature series were used for each of the 16 California Title 24 Building Climate Zones.[2] Title 24 divides California into 16 climate zones based on similarity of temperature, typical energy consumption, and other factors. The list of reference cities is shown in Table 4-19.

---

2 California Energy Commission. *Climate Zone Weather Data Analysis and Revision Project*. Augustyn and Company, March 22, 1991.

Table 4-19: T24 Climate Zones

T24 Zone	T24 Zone Reference City
Zone 1	Arcata
Zone 2	Santa Rosa
Zone 3	Oakland
Zone 4	Sunnyvale
Zone 5	Santa Maria
Zone 6	Los Angeles
Zone 7	San Diego
Zone 8	El Toro
Zone 9	Pasadena
Zone10	Riverside
Zone11	Red Bluff
Zone12	Sacramento
Zone13	Fresno
Zone14	China Lake
Zone15	El Centro
Zone16	Mount Shasta



*California Energy Commission Cartography Unit

Downloaded from http://www.energy.ca.gov/maps/building_climate_zones.html

Households were mapped to the T24 zone using the service address information. Average weather for each zone was based on a reference city. Two sources of weather data were used in this analysis.

- *Actual daily average dry-bulb temperatures* – For every day during the period for which utility billing data were provided, daily average temperatures were obtained from NOAA for the weather stations nearest the cities in Table 4-19.
- *Normal-year temperatures* – Eleven years (1999 through 2009) of daily average drybulb temperatures were averaged, by Julian date, to create a normal weather year for each of the sixteen weather stations.

Degree-Day Normalization Methodology

The DDN methodology modeled monthly consumption as a function of monthly heating degree days and cooling degree days (HDD and CDD, respectively). The HDDs and CDDs for each household reflected the sum of daily degree-day series. Heating degree-days for a particular day was the difference between the heating degree-day base τ_1 and the daily average temperature, if the daily average was below the base, and 0 if the daily average was above the base. Similarly, cooling degree-days for the day was the difference between the daily average temperature and the cooling degree-day base τ_2, if the daily average was above the base, and 0 if the daily average was below the base. The base or reference temperatures τ_1 and τ_2 were specific to each household, based on the model fit.

This relationship is shown in Equation 4-1. For each unique billing series, the coefficients β_0, β_1, and β_2, and the parameters τ_1 and τ_2 were estimated so as to best fit the relationship between outdoor temperature and monthly energy consumption.

Equation 4-1: The DDN Heating and Cooling Model

$$U_i = \beta_0 + \beta_1 * HDD_i(\tau_1, T_{ext}) + \beta_2 * CDD_i(\tau_2, T_{ext}) + e_i$$

$$HDD_i(\tau_1, T_{ext}) = \sum_{d=firstdayi}^{lastdayi} \max(\tau_1 - T_{ext\,d}, 0)$$

$$CDD_i(\tau_2, T_{ext}) = \sum_{d=firstdayi}^{lastdayi} \max(T_{ext\,d} - \tau_2, 0)$$

Where:

U_i = Electric usage during billing cycle i.

T_{ext} = Series of external temperatures for each day of the study period

Text d = External temperatures on day d

HDD_i = Sum of heating degree days based on reference temperature τ_1 during billing cycle i.

CDD_i = Sum of cooling degree days based on reference temperature τ_2 during billing cycle i.

β_0 = Estimate of the average daily base load (temperature-invariant component of usage)

β_1 = Increase in electric usage for each incremental increase in heating degree days

β_2 = Increase in usage for each incremental increase in cooling degree days

τ_1 = Outside temperature at which the household's heating-related usage begins

τ_2 = Outside temperature at which the household's cooling-related usage begins

e_i = Residual Error

This normalization procedure did not include other weather phenomena that impact energy consumption, such as humidity and rainfall, or insulation and daylight. This was to ensure that the number of unique billing periods was substantially larger than the number of variables in the model, to avoid over-fitting.

Roughly speaking, the term β_1HDD corresponded to heating and the term β_2CDD corresponded to cooling. However, non-heating and non-cooling uses also varied over the year in ways that were correlated to some extent with heating and cooling degree-days. Thus, these estimated terms also included both positive and negative seasonal effects associated with other uses.

Therefore, heating and cooling use were not assumed to be given by these terms. Instead, heating and cooling were estimated via the cross-sectional CDA analysis applied to the total Normalized Annual Consumption. Heating and cooling coefficients were used as indicators of the presence of heating and cooling, as described in the CDA section later in this chapter.

For some households, one or both of the degree-day terms showed little relationship to monthly consumption. The analysis used an F-test as a diagnostic to determine for a particular household whether to include either the heating or cooling term, both terms, or neither term. Also, the heating or cooling term was dropped from the model if its coefficient was negative.

For electricity, the analysis tested for inclusion of both heating and cooling terms. Based on the diagnostics, the best-fit model for a particular household included both heating and cooling terms, only a heating term, only a cooling term, or neither of the terms. Inclusion of neither term meant that only a base term β_0 was estimated. For gas, the analysis assumed no gas cooling. [3] The gas model for a household therefore included only heating, or only a base term.

For each set of reference temperatures, the normal-year HDD and CDD were calculated at all weather stations. The appropriate normal-year HDD and CDD series for each household were applied to the household's estimated coefficients from the DDN model to provide the predicted Normalized Annual Consumption (NAC), which formed the basis for the subsequent end-use analysis. The next section discusses the DDN models for the survey respondents, while the following section discusses the normalized consumptions that were predicted for those households.

Electric DDN Models

The distribution of households with electric and gas DDN models by climate zone and utility is shown in Tables 4-20 and 4-21. Only households with insufficient or unrealistic billing series did not have DDN models.

---

3 The CDD term is empirically small, and there is no gas cooling in single-family dwellings. Where there is no strong cool-weather-related trend, the CDD regression coefficient will reflect any deviations from the best-fit equation, essentially fitting to noise.

Table 4-20: Number of Households With Electric DDN Models by T24 Zone and Utility

T24 Zone	Number of Households Modeled				
	SDG&E	LADWP	PG&E	SCE	All Utilities
1 (Arcata)	.	.	132	.	132
2 (Santa Rosa)	.	.	689	.	689
3 (SF/ Oakland/ Monterey)	.	.	2,241	.	2,241
4 (Santa Jose)	.	.	971	1	972
5 (Santa Maria)	.	.	238	152	390
6 (Los Angeles)	236	340	.	2,027	2,603
7 (San Diego Lindbergh)	2,486	.	.	.	2,486
8 (Santa Ana El Torro)	87	224	.	1,889	2,200
9 (Burbank)	.	1,465	.	1,636	3,101
10 (SD/ Mirmr/ Riverside)	1,044	.	.	1,500	2,544
11 (Red Bluff)	.	.	626	.	626
12 (Sacramento)	.	.	1,588	.	1,588
13 (Fresno)	.	.	767	762	1,529
14 (China Lake)	36	.	.	794	830
15 (El Centro)	16	.	.	924	940
16 (Mt. Shasta/ Bishop)	.	.	162	911	1,073
All California	3,905	2,029	7,414	10,596	23,944

Source: *2010 California Residential Appliance Saturation Survey*

Table 4-21: Number of Households with Gas DDN Models by T24 Zone and Utility

T24 Zone	Number of Households Modeled			
	PG&E	SoCal Gas	SDG&E	All Utilities
1 (Arcata)	83	.	.	83
2 (Santa Rosa)	496	.	.	496
3 (Sf/ Oakl/ Montry)	1,852	.	.	1,852
4 (Santa Jose)	776	.	.	776
5 (Santa Maria)	182	118	.	300
6 (Los Angeles)	.	1,492	115	1,607
7 (San Diego Lindbergh)	.	.	2,028	2,028
8 (Santa Ana El Torro)	.	1,801	42	1,843
9 (Burbank)	.	3,009	.	3,009
10 (Sd/ Mirmr/ Rvrsde)	.	1,291	757	2,048
11 (Red Bluff)	402	.	.	402
12 (Sacramento)	1,210	.	.	1,210
13 (Fresno)	644	558	.	1,202
14 (China Lake)	.	282	.	282
15 (El Centro)	.	623	.	623
16 (Mt. Shst/ Bishop)	13	358	.	371
All California	5,658	9,532	2,942	18,132

Source: *2010 California Residential Appliance Saturation Survey*

Table 4-22 shows how the type of best-fit model for electric varied by T24 zone. For Zones 1 through 7, along the coast of California, consumption was found to be either constant over the year (and fit a base load only), or responsive to heating degree days but not cooling. In Zones 8-10, along the Southern California coast, households either fit a base load model or were responsive to cooling degree-days only. In the remaining zones, most households were responsive to CDD (with and without also responding to HDD).

Table 4-22 Best Fit DDN Electric Model Type by T24 Zone

T24 Zone	Electric Model Type								All California
	Base Load Only		Heating and Base Load		Cooling and Base Load		Heating, Cooling and Base		
	#Resp	%Resp	#Resp	%Resp	#Resp	%Resp	#Resp	%Resp	#Resp
1 (Arcata)	66	50.0%	66	50.0%	.	.	.	.	132
2 (Santa Rosa)	337	48.9%	306	44.4%	29	4.2%	17	2.4%	689
3 (SF/ Oakland/ Monterey)	1,044	46.5%	1,143	51.0%	29	1.2%	25	1.1%	2,241
4 (Santa Jose)	439	45.1%	359	36.9%	129	13.2%	45	4.6%	972
5 (Santa Maria)	211	54.1%	170	43.5%	4	1.0%	5	1.2%	390
6 (Los Angeles)	1,230	47.2%	1,021	39.2%	268	10.2%	84	3.2%	2,603
7 (San Diego Lindbergh)	1,614	64.9%	602	24.2%	200	8.0%	70	2.8%	2,486
8 (Santa Ana El Torro)	1,039	47.2%	363	16.5%	733	33.3%	65	2.9%	2,200
9 (Burbank)	940	30.3%	324	10.4%	1,644	53.0%	193	6.2%	3,101
10 (SD/ Mirmr/ Riverside)	476	18.7%	191	7.5%	1,688	66.3%	189	7.4%	2,544
11 (Red Bluff)	100	15.9%	43	6.8%	380	60.7%	103	16.4%	626
12 (Sacramento)	347	21.8%	274	17.2%	723	45.5%	244	15.3%	1,588
13 (Fresno)	128	8.3%	49	3.2%	1,217	79.5%	135	8.8%	1,529
14 (China Lake)	160	19.2%	60	7.2%	519	62.5%	91	10.9%	830
15 (El Centro)	204	21.7%	37	3.9%	672	71.4%	27	2.8%	940
16 (Mt. Shasta/ Bishop)	356	33.1%	225	20.9%	420	39.1%	72	6.7%	1,073
All California	8,691	36.2%	5,233	21.8%	8,655	36.1%	1,365	5.7%	23,944

Source: *2010 California Residential Appliance Saturation Survey*

The model calculated increasing consumption with colder weather below the HDD set-point and increasing consumption with hotter weather above the CDD set-point. Some variation across the T24 zones was found, as shown in Table 4-23. There was generally good agreement between the HDD set-point for households with and without a CDD set-point; similarly, CDD set-points agreed between households with and without an HDD set-point. For California as a whole, the average set-point for cold weather was 58° Fahrenheit[4] (F) while the average hot weather set-point was 75.8°F. This demonstrated the value of choosing the best set-points for each household rather than using 65°F for both HDD and CDD.

Table 4-23: Average Degree-Day Reference Temperature (°F) by Electric Best-Fit Model Type

T24 Zone	Electric Model Type			
	Heating and Base Load (N=5,271)	Cooling and Base Load (N=8,525)	Heating, Cooling and Base Load (N=1,331)	
	HDD Ref.	CDD Ref.	HDD Ref.	CDD Ref.
1 (Arcata)	56.8°			
2 (Santa Rosa)	57.8°	72.9°	54.3°	73.8°
3 (Sf/ Oakland/ Monterey)	56.2°	74.8°	55.4°	74.7°
4 (Santa Jose)	56.7°	75.3°	56.4°	75.7°
5 (Santa Maria)	54.2°	74.7°	50.4°	75.2°
6 (Los Angeles)	59.5°	74.0°	58.4°	75.9°
7 (San Diego Lindbergh)	59.5°	74.0°	57.6°	75.5°
8 (Santa Ana El Torro)	57.4°	73.2°	57.0°	74.2°
9 (Burbank)	56.0°	73.0°	55.2°	74.4°
10 (SD/ Mirmr/ Riverside)	59.4°	74.6°	55.8°	77.5°
11 (Red Bluff)	59.2°	75.6°	58.5°	76.3°
12 (Sacramento)	58.6°	73.6°	57.4°	74.1°
13 (Fresno)	57.8°	74.6°	58.8°	74.9°
14 (China Lake)	60.4°	76.3°	58.0°	76.0°
15 (El Centro)	72.0°	78.0°	64.1°	77.2°
16 (Mt. Shasta/ Bishop)	56.5°	73.7°	53.8°	74.5°
All California	57.8°	74.4°	57.0°	75.3°

Source: *2010 California Residential Appliance Saturation Survey*

Gas DDN Models

As shown in Table 4-24, the majority of best-fit DDN models for gas were heating plus base load models. This was expected given that most residential gas was used for household heating and water heating. Eighty-eight percent of households with gas had a cold-weather dependent term (heating and base load model) while the remainder (12%) predicted constant consumption regardless of outdoor temperature (base load only).

4 All temperatures are reported in degrees Fahrenheit.

Heating reference temperatures were on average 61.8°F, with some variation by T24 zone. This reference is between four and five degrees higher than the reference temperature of the electric consumption models.

Table 4-24: Best-Fit DDN Gas Model Type by T24 Zone

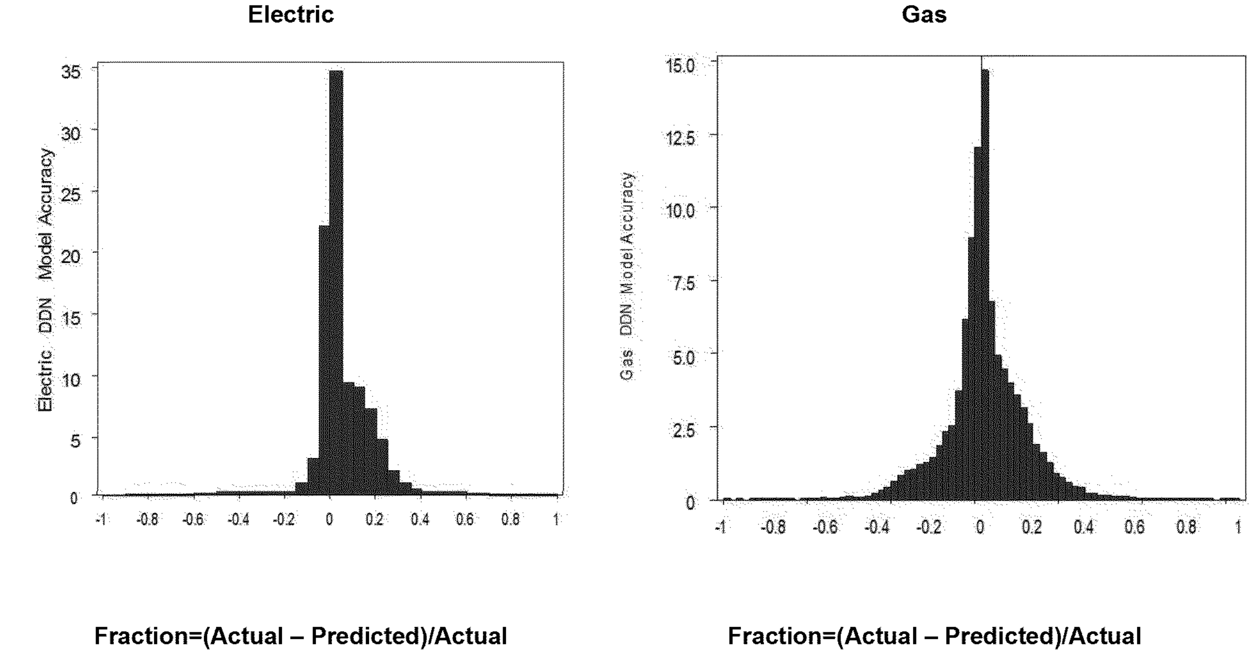
T24 Zone	Gas Model Type				
	Base Load Only		Heating and Base Load		
	#Resp	%Resp	#Resp	%Resp	HDD Ref Temp °F
1 (Arcata)	4	5%	79	95%	60.7°
2 (Santa Rosa)	23	5%	473	95%	59.2°
3 (SF/ Oakland/ Monterey)	215	12%	1,637	88%	58.0°
4 (Santa Jose)	51	7%	725	93%	59.8°
5 (Santa Maria)	34	11%	266	89%	56.1°
6 (Los Angeles)	216	13%	1,391	87%	62.2°
7 (San Diego Lindbergh)	277	14%	1,751	86%	62.8°
8 (Santa Ana El Torro)	292	16%	1,551	84%	61.2°
9 (Burbank)	463	15%	2,546	85%	62.0°
10 (SD/ Mirmr/ Riverside)	193	9%	1,855	91%	64.5°
11 (Red Bluff)	17	4%	385	96%	63.3°
12 (Sacramento)	62	5%	1,148	95%	61.5°
13 (Fresno)	73	6%	1,129	94%	62.8°
14 (China Lake)	35	12%	247	88%	64.7°
15 (El Centro)	143	23%	480	77%	67.6°
16 (Mt. Shasta/ Bishop)	45	12%	326	88%	57.9°
All California	2,143	12%	15,989	88%	61.8°

Source: *2010 California Residential Appliance Saturation Survey*

NAC Results

The Normalized Annual Consumption estimates derived from the electric and gas DDN models were generally close to the actual annualized consumption. Electric DDN models predicted about 5.7% less than the actual consumption for the sample. Gas DDN models predicted about 1% less gas consumption than was actually billed. These differences reflect that the normal temperature series were slightly but not dramatically different from the actual temperatures.

Figure 4-2: Comparison of NAC Results and Actual Consumption



Source: *2010 California Residential Appliance Saturation Survey*

CDA Variables and Data Imputation Process

This section addresses additional treatment of survey responses required for the conditional demand analysis (CDA) to produce unbiased unit energy consumption (UEC) estimates. The CDA was restricted to individually metered accounts; therefore, the following discussion applies only to individually metered survey responses. The following processes are covered in this section.

- Creation of binary variables indicating the presence of a particular end use of the fuel being modeled.
- Creation of continuous variables to reflect intensity of system use.
- Data imputation processes for missing values.

In this section, variables that received similar processing are grouped together. The first group includes four variables: square footage, age of dwelling, number of residents, and household income. These four variables were critical to the all UEC estimates, and therefore, special attention was given to the process used to impute missing values. Next, the space heating and water heating variables are discussed with special attention given to the process used to identify fuel misreporting. The creation of indicator and continuous variables is also discussed. Finally, the refrigerators and freezers are discussed with particular attention to initial engineering estimates for energy use of these appliances.

SB_GT&S_0771254

Advanced Variable Imputation Process

The survey variables for square footage, household income, age of the dwelling, and number of residents were critical to the development of multiple UEC estimates. Because of their importance to the analysis, for these four variables it was essential to minimize bias that may result from imputing missing values. The missing values were imputed using a regression-based approach as described in the following steps:

1. A binary variable was created for each of the four variables that took the value of 1 if the respondent answered the question and 0 if not.
2. The indicator variables served as dependent variables in a logistic regression used to estimate the likelihood of response to the specific question. Other survey responses served as the independent variables to the logistic regression.
3. The probability of response to each question was calculated based on the logistic regression and used to estimate an inverse Mills' ratio for each respondent for each of the four questions.
4. A linear regression model was estimated to provide a predicted response value for those who did not answer the question. The inverse Mills' ratio was included in the linear regression as an explanatory variable, controlling for non-response bias.

If a survey was missing any of the explanatory variables needed to apply the regression-based approach, the missing values were imputed using a conditional means process, which involved calculating the mean value by dwelling type.

Square Footage and Surface Area

The survey collected data on square footage in the SQFT variable for a series of size ranges. To use the size ranges in the CDA, they had to be converted to a continuous series, which was recorded in the *SQFT_A* variable. This variable typically used the mid-point of each size range to provide an estimate of the dwelling's square footage. However, for responses in three of the size ranges the following sizes were assumed:

- Dwellings in the smallest size group, less than 250 square feet, were assumed to be 200 square feet.
- Dwellings that were between 4,001 – 5,000 square feet were assumed to be 4,700 square feet.
- Dwellings greater than 5000 square feet were assumed to be 6,000 square feet.

If *SQFT* was missing, the value was imputed using the means value by dwelling type.

Once the continuous square foot series was assigned to each respondent, these estimates were converted to a new variable, AREA, which provided an estimate of the dwelling's surface area. This was calculated using the following equations from the 2003 RASS[5] presented below.

For single-story, single-family dwellings and mobile homes

$$surface\ area = 5.9985 * SQFT\_A^{0.8528}$$

For multi story, single-family dwellings

$$surface\ area = 13.9694 * SQFT\_A^{0.7395}$$

For multi-family dwellings

$$surface\ area = 0.5955 * SQFT\_A^{1.1034}$$

Missing values for the SQFT_A variable were imputed using the regression-based approach outlined above.

Household Income

The survey collected data on household income in the INCOME variable for a series of income ranges. The INCOME variable was converted into a continuous variable AVGINC, which was the mid-point of each of the income ranges, except for the highest income group of $150,000 or more. Responses in this income range were set to $175,000. If *INCOME* was missing, *AVGINC* was imputed using the mean value by dwelling type.

Dwelling Age

The survey collected data on the year a dwelling was built in the *BUILTYR* variable, to which respondents selected from a series of age ranges. The *BUILTYR* variable was converted into a continuous *BUILT* variable, which was the mid-point of each of the age ranges, except for the oldest and most recent age ranges. Responses in these categories were assigned values of 1935 and 2007, respectively.

During the Data Imputation, a new variable *HOMEAGE* was created that contained the value of *BUILT* or an imputed value using the regression-based approach discussed above. If no regression value could be derived, then *HOMEAGE* was imputed with the average age by dwelling type. In addition, the binary variable *NEWHOME* was set equal to 1 for all dwellings built after 2000, and zero otherwise.

Number of Household Residents

The survey collected data on the number of residents by age group in variables for each age group. These responses were summed to create a count of the total number of people in the

---

5 *California Statewide Residential Appliance Saturation Study Final Report*. June 2004. CEC 400-04-009.

household, *RESCNT*. *RESCNT* was required for the CDA so missing values were imputed using the regression-based approach to create a new variable (*NUMI*) set equal to *RESCNT* or the imputed value when *RESCNT* was missing. The CDA required a log transformation of the *NUMI* variable so a new variable was created (*NHH*) and set equal to (log (*NUMI* + 1)).

Correcting for Fuel Misreporting

Previous CDA studies conducted on the California residential population have shown that misreporting of fuels used for heating and water heating was common, particularly in multi-family units and in areas with very low consumption. The variables used to model each observation must reflect an accurate profile of each observation for the statistical technique used for this CDA to provide accurate results. Since space and water heat account for such a large share of an individual's energy consumption, it is critical to identify cases in which a respondent inaccurately misreported whether they have electric or gas space and water heat.

After the fuel switching validation process was completed, binary and continuous variables were derived from survey responses to identify the presence of electric and gas systems and the degree to which systems were used. These variables are also covered below.

Space Heating

The space heating section gathered information for both primary and secondary heating systems, as well as temperature settings that indicate intensity of use. Survey responses concerning an individual's primary and secondary space heating systems were cross referenced with the following information to determine whether survey responses were accurate:

Primary Heat

Natural Gas Line – As discussed above, the variable NGLINE2 recorded whether a residence was found to have natural gas, by cross-referencing survey and utility information. For respondents who reported having primary electric heat, the presence of a gas line provides evidence that they may actually have gas heat.

Utility Heat Allowance – Electric utilities offer a separate tariff that provides an allowance for households with electric heat. The tariff generally allows electric heat households to pay a lower rate per KWH than non-heating households. Each of the utilities provided information regarding which respondents received an electric heating allowance.

Significant Coefficient on HDD From the Electric DDN – If a household's electric consumption was responsive to an increase in heating degree days, then the electric DDN model for that household was likely to have a significant coefficient on HDD. While those without a heating term may have electric heat that they do not use, the presence of a heating coefficient in the electric DDN model was considered an indication of fuel misreporting.

The specific rules used to determine fuel misreporting are outlined below. The rules only apply to households for which the variable indicating that they pay for heat (*PAYHEAT2*) was "yes." The fuel misreporting rules did not apply to households that did not pay for heat because survey respondents only asked to fill out the heating section if they paid for heat. The variable *PHTFUEL2* recorded the imputed primary heat fuel.

The following rules were applied to households that reported having primary electric heat and were identified as having gas in the residence:

- If there was not a significant heating parameter in the electric DDN model, then PHTFUEL2 = 1 (gas)
- If the utility does not identify them as having electric heat, then PHTFUEL2 = 1 (gas).
- If there was a significant heating term in the electric DDN model and the utility indicated they have electric heat, then PHTFUEL2 = 2 (electric). A flag was constructed for these respondents to determine whether their CDA parameter estimates differed from other respondents with electric heat. Therefore, the only way a household with a gas line was allowed to have electric heat was if both the DDN model and the electric utility showed evidence of electric heat.

The following rules were applied to people who reported having primary electric heat who were identified as not having a gas line to the residence based on the cleaning process outlined in earlier in this chapter.

- If there was not a significant heating parameter in the electric DDN model, then PHTFUEL2 = 6 (other)
- If there was a significant heating parameter in the electric DDN model, then PHTFUEL2 = 2 (electric). A flag was constructed for cases in which the utility did not provide indication of electric heat to determine whether these CDA parameter estimates differed from other respondents with electric heat.

Table 4-25 shows the distribution of respondents according to reported *PHTFUEL* and cleaned *PHTFUEL2*.

Table 4-25: Primary Space Heating Fuel Cleaning

PHTFUEL	*PHTFUEL2*							Total	% Total
	Natural Gas	Electric	Bottled Gas	Wood	Other	Missing	N/A		
Natural Gas	17,197	0	0	0	183	0	0	17,380	71.0%
Electric	1,632	1,479	0	0	869	0	0	3,980	16.3%
Bottled Gas	0	0	847	0	0	0	0	847	3.5%
Wood	0	0	0	560	1	0	0	561	2.3%
Other	0	0	0	0	43	0	0	43	0.2%
Missing	3	0	1	0	10	715	2	728	3.0%
N/A	0	0	0	0	0	0	922	922	3.8%
All	18,832	1,479	848	560	1,106	715	924	24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

Auxiliary Heat

Respondents were only allowed to have alternative electric heat if one of the following was true:

- The alternative electric heating was baseboard heating.
- The alternative electric heating was a heat pump with primary wood stove or fireplace.
- The alternative electric heating was a portable heater.
- For these cases, the imputed alternative heat variable (AHFUEL2) was set to 2 (electric), while all other cases it was set to 6 (other).

Table 4-26 shows the distribution of respondents according to *AHTFUEL* and *AHTFUEL2*.

Table 4-26: Auxiliary Space Heating Fuel Cleaning

AHTFUEL	*AHTFUEL2*								Total
	Natural Gas	Electric	Bottled Gas	Wood	Solar	Other	Missing	N/A	
Natural Gas	3,033	0	0	0	0	0	0	0	3,033
Electric	0	2,797	0	0	0	5,606	0	0	8,403
Bottled Gas	0	0	210	0	0	0	0	0	210
Wood	0	0	0	1,192	0	0	0	0	1,192
Solar	0	0	0	0	25	0	0	0	25
Other	0	0	0	0	0	19	0	0	19
Missing	0	0	0	0	0	0	368	0	368
N/A	0	0	0	0	0	0	0	11,214	11,214
All	3,033	2,797	210	1,192	25	5,625	368	11,214	24,464

Source: *2010 California Residential Appliance Saturation Survey*

Space Heating Binary Variables

The following indicator variables were also derived from survey responses to reflect the presence or absence of each respective space heating technology:

Electric Heat

- DEHEAT –If the household pays for heat and the primary heating fuel is electric, DEHEAT was set equal to one, zero otherwise.
- NONELEBK – If the household has a primary electric heater and a non-electric backup, NONELEBK was set equal to one, zero otherwise.
- DEAUXHT – If the household has an additional electric heater, DEAUXHT was set equal to one, zero otherwise.
- ROOM – If the household has electric heat and the primary heater is a resistance heater, a through the wall heat pump, or a portable heater, ROOM was set equal to one, zero otherwise.

Gas Heat

- DGHEAT – If the household pays for heat and the primary heating fuel is natural gas, DGHEAT was set equal to one, zero otherwise.
- NONGBU – If the household has a primary gas heater and a non-gas backup, NONGBU was set equal to one, zero otherwise.
- DNGAUXHT – If the household has an additional natural gas heater, DNGAUXHT was set equal to one, zero otherwise.
- GROOM – If the household has gas heat and the primary heater is a floor or wall furnace, GROOM was set equal to one, zero otherwise.
- SETBK – An indicator variable accounting for people who lower the thermostat setting at night. SETBK was set equal to one for respondents whose nighttime heater setting (HNITESET) was lower than the average setting (HTTSET), otherwise SETBK was set equal to zero.

Space Heating Continuous Variable

- HTTSET – The average daily thermostat temperature was set equal to the weighted average of each household's thermostat temperature for each time period during the heating season. Missing values for this variable were imputed with the mean value by dwelling type.

Primary Water Heat

Survey responses concerning a household's primary water heater were cross referenced with billing information to determine whether survey responses were accurate. The following steps were used to evaluate whether respondents that indicated they had an electric hot water heater actually had a gas hot water heater.

- Summer months were identified as the three warmest months of the year by climate zones.
- Any respondent that indicated having an electric hot water tank and also had natural gas in the home were identified.
- The average monthly gas consumption over the three warmest months of the year for those households was calculated.
- Households with more than ten therms per month over the summer were flagged by setting GWH_FLAG =1.
- When GWH_FLAG =1, the new variable PRWHFUEL2 was set to 1 (gas), and for all other households, PRWHFUEL2 was set = PRWHFUEL.

Table 4-27 shows a comparison of *PRWHFUEL* and *PRWHFUEL2*.

Table 4-27: Primary Water Heating Fuel Cleaning

PRWHFUEL	*PRWHFUEL2*							Total	% Total
	Natural Gas	Electric	Bottled Gas	Solar	Other	Missing	N/A		
Natural Gas	17,618	0	0	0	0	0	0	17,618	72.0%
Electric	414	1,961	0	0	0	0	0	2,375	9.7%
Bottled Gas	23	0	1,112	0	0	0	0	1,135	4.6%
Solar	2	0	0	7	0	0	0	9	0.0%
Other	2	0	0	0	16	0	0	18	0.1%
Missing	364	0	0	0	0	618	0	982	4.0%
N/A	655	0	0	0	0	0	1,672	2,327	9.5%
All	19,078	1,961	1,112	7	16	618	1,672	24,464	100.0%

Source: *2010 California Residential Appliance Saturation Survey*

Water Heating Binary Variables

The following indicator variables were also derived from survey responses to reflect the presence or absence of electric or natural gas water heating:

The CDA contains a gas, an electric, and a solar water heating fuel indicator variable.

Electric Water Heating

- DEWH * Set equal to one for respondents with a electric hot water heater, zero otherwise.
- DWHSOLAR * Set equal to one for respondents with a solar hot water heater with an electric backup, zero otherwise.
- ADDWHEL * Set equal to one for respondents with more than one electric water heater, zero otherwise.

Gas Water Heating

- DGWH * Set equal to one for respondents with a gas hot water heater, zero otherwise.
- DGWHSOLAR * Set equal to one for respondents with a solar hot water heater with a natural gas backup, zero otherwise.

Water Heating Continuous Variables

The primary drivers of energy consumption for water heaters are clothes washers, dishwashers, and showers or baths. The following continuous variables were constructed to account for hot water usage due to these appliances:

- CWASHU * Clothes washer usage constructed from the number of loads per day by water temperature.

- DWASHU – Dishwasher usage constructed from number of loads per day.
- WHTSHWRS – Total number of baths and showers taken per day.

For respondents that did not answer the usage questions, *DWASHU*, *CWASHU,* or WHTSHRS were imputed using the mean value by dwelling type.

Binary and Continuous CDA Variables

This section reports on a number of survey variables for which both continuous and binary variables were constructed. The binary variables reported on the presence of each respective appliance while the continuous variables provided an indication of the amount or intensity of appliance use.

Central Air Conditioning

- DCAC – Set equal to one to indicate the presence of a central air conditioner, zero otherwise.
- TSETC – Continuous variable for the weighted average of the thermostat temperature for each time period during the cooling season. If the household had central air conditioning and did not report the temperature, the mean value by dwelling type was assigned.

Room Air Conditioning

- DRAC – Set equal to one to indicate the presence of room air conditioning, zero otherwise.
- RACCNT – Count of the number of room air conditioners.
- TSETUSE – Continuous variable for the weighted average of the room air conditioner use. If the household had a room air conditioner and did not report the temperature, the mean value by dwelling type was assigned.

Clothes Dryers

CDA variables for clothes dryers included the following indicator variables reporting the presence of electric or gas dryers and two continuous variables reporting the number of loads per week.

- GDRY – Set equal to one to indicate the presence of a gas clothes dryer that was not in a common area, zero otherwise.
- EDRY – Set equal to one to indicate the presence of an electric clothes dryer that was not in a common area, zero otherwise.

- GDRYU – The weekly usage of the gas dryer. If the survey response to DRYLDS was missing, and the household had a gas dryer, GDRYU was imputed using the mean value by dwelling type.
- EDRYU – The weekly usage of the electric dryer. If the survey response to DRYLDS was missing, and the household had an electric dryer, EDRYU was imputed using the mean value by dwelling type.

Outdoor Lighting

CDA variables for outdoor lighting included the following indicator and continuous variables.

- DOLT – Set equal to one to indicate the presence of exterior lighting, zero otherwise.
- OLTFIX – Total number of exterior fixtures.
- OPROPHID – Continuous variable for the proportion of exterior lighting fixtures that were HID lights.
- OPROPSEN – Continuous variable for the proportion of exterior lighting fixtures that were on sensors.
- OPROPTIM – Continuous variable for the proportion of exterior lighting fixtures that were on timers.

Televisions

CDA variables for televisions included the following indicator and continuous variables.

- DTV – Set equal to one to indicate the presence of either standard, small LCD, large LCD or plasma TV, zero otherwise.
- TVHRS – The sum of the total number of hours watching small and large screen TVs per day.
- TVKW – Variable that accounts for electricity use based upon number of hours of usage. Standard and small LCD TVs were assumed to use 0.1 kWh per hour, and large screen LCD and plasma TVs were assumed to use 0.25 kWh per hour. If the household had multiple types of TVs, the usage numbers were multiplied by the proportion of TVs of each type.

If the household had one or more televisions and usage information was missing, the mean value by dwelling type was assigned.

Personal Computers and Home Offices

The current RASS collected data on the number desktop and laptop PCs as well as the number of hours each was used. These variables were converted to the same variables used in the 2003 study to maintain continuity. For each

- DPC – Set equal to one to indicate the presence of either a desktop or laptop personal computer, zero otherwise.

- PCHRS – Continuous variable for the sum of desktop and laptop PC hours.
- PCNUM – The sum of the number of hours of usage for desktop and laptop PCs.
- DHMOFF – Set equal to one to indicate that someone in the household operated a business or worked from home, zero otherwise.
- HMOFFHRS – Continuous variable for the numbers of hours a week someone works out of the home.

Where applicable, for any respondent who did not provide a response to any of the three continuous variables, PCHRS, PCNUM, and HMOFFHRS, the value was imputed with the mean value by dwelling type.

Swimming Pools

CDA variables for pools included the following indicator and continuous variables.

- DPLPMP – Set equal to one to indicate the presence of a pool if the respondent indicated that they pay for its energy use. Only single family households were allowed to have pools. All other pools listed in the survey were assumed to be pools located in common areas, and were disallowed in the CDA.[6]
- PLFILT – The number of hours per day used to filter the pool. This variable differs between summer months (May-October) and winter months (November-April).
- PLSIZE – The pool size variable was set to 18,000 gallons for small pools, 30,000 for medium sized pools, and 42,000 for large pools.
- EPLHT – Set equal to one to indicate that the pool was heated with electric heat, zero otherwise.
- DGPLHT – Set equal to one to indicate that the pool was heated with natural gas, zero otherwise.
- GPLHTFREQ – The gas CDA also analyzed the impact of the frequency of pool heating. This variable was allowed to differ between summer and winter months.
- PLCOV – Set equal to one to indicate the use of a pool cover. A pool cover may reduce the heating needs due to an increase in pool temperature or it may indicate a pool that is used more frequently, leading to an increase in heating needs.

Spas and Hot Tubs

The CDA required the following indicator variables regarding the presence of a spa or hot tub and the fuel type.

---

6 Respondents could have answered yes, "I have a pool and I pay for its energy use," when pools were located in common areas. Home owners' association fees often include a set amount for the expense of heating and filtering common area pools. To help reduce this possibility, we restricted the analysis of pools to single-family dwellings.

- DSPA – Set equal to one if the respondent lived in a single family dwelling, townhouse, or mobile home, had a spa or hot tub, and paid for its energy use, was set equal to one, zero otherwise.
- DEHTSPA – Set equal to one for spas heated with electricity or solar with electric backup, zero otherwise
- SPASOLAR – Set equal to one for spas heated by solar with electric backup, zero otherwise.
- DGHTSPA – Set equal to one for spas heated with natural gas or solar with natural gas backup, zero otherwise.
- SPAGSOLAR – Set equal to one for spas heated by solar with natural gas backup, zero otherwise.

The CDA also required the following continuous variables regarding spa filter and heat usage by fuel type and spa size.

- SPCOV – Set equal to one if the spa had an insulated cover, zero otherwise.
- SPAFREQ – The frequency of spa filtering.
- SPAEHTFREQ – The frequency of electric heating was allowed to differ between summer and winter months.
- SPAGHTFREQ – The frequency of natural gas heating was allowed to differ between summer and winter months.
- SPASIZE – Continuous variable based on the number of people the spa holds. The number of people was set to 2 for small spas, 5 for medium spas, and 8 for large spas.

Fans

The CDA incorporated three types of fans: forced air fans, attic fans, and ceiling fans.

- DFFAN – Set equal to one if the primary heating fuel was natural gas or bottled gas and the heater is a central heater, zero otherwise.
- DATTFAN – Set equal to one if the household has an attic or a whole house fan, zero otherwise.
- DCEILF – Set equal to one if the household has at least one ceiling fan, zero otherwise.

Seasonal Home Indicator

The CDA accounted for differences in energy consumption between year-round and seasonal homes using the following variable:

- SEASONAL – Set equal to one for anyone that reported the residence was not their year-round residence and lived there less than 12 months of the year, zero otherwise.

Double-Pane Windows and Dwellings in Colder Zones

- DPWIN – Set equal to one for respondents who indicated their WINDTYPE was all or mostly double paned or a mixture of double and single, zero otherwise.
- T24 CZ – Set equal to one if the residence was in zone 1 or 16 and zero otherwise. Title 24 has building requirements that apply to new dwellings in CEUS weather zones 1, 161 and 162, which may offset the colder climates in these zones.

Kitchen Appliances

The following indicator variables were defined for kitchen appliances:

- DERGOV – Set equal to one for households with either an electric range or oven, zero otherwise.
- DGRGOV – Set equal to one for households with either a natural gas range or oven, zero otherwise.
- DMWV – Set equal to one for households that indicated they had a microwave oven, zero otherwise.

Laundry

- DCW – Set equal to one to indicate the presence of laundry equipment in the home and either a top loading or a front-loading washer, zero otherwise.

Energy Consumption for Refrigerators and Freezers

Engineering estimates used in the CDA model accounted for differences in energy consumption of refrigerators and freezers with differing characteristics. This step was necessary because these appliances had roughly 100% saturation, eliminating differences among households with and without refrigerators and freezers. Without such differences, the statistical model used to estimate the UECs would be unable to identify the energy consumption of the appliances. Using pre-defined engineering estimates for refrigerators and freezers with different characteristics, the model was better able to detect variation in consumption between households with different refrigerators and freezers.

As was previously mentioned, the process used for the current CDA sought to maintain consistency with the 2003 RASS. Therefore, engineering estimates in the current study were based on those used in the 2003 RASS. Those estimates were based on data provided by the Association of Home Appliance Manufacturers (AHAM) website (www.aham.org). Refrigerators were tied to the AHAM data based on door style, size, defrost, and age characteristics. Freezers were tied to the AHAM data based on style, size, and age characteristics. Each of these characteristics was the same between the two surveys except age. Ages from the 2009 survey were grouped to tie the current survey responses to the 2003 engineering estimates. Refrigerators that were less than two years old in the 2009 survey were given the AHAM consumption estimate for a similar style and size refrigerator as used in the 2003 RASS. Refrigerators between 2 and 10 years old were given the energy of 2 to 7 year old refrigerators from the 2003 survey. The remaining refrigerators were assigned the same energy use as 8 to 10 year old refrigerators from the 2003 RASS.

CHAPTER 5: Data Analysis Methodology

The conditional demand analysis (CDA) used to derive electric and natural gas unit energy consumption (UEC) estimates employed a statistically adjusted engineering (SAE) analysis modeling technique. The SAE model implemented was similar to that used in the 2006 Update[7] to the 2003 RASS, in which engineering estimates were based on the 2003 RASS equations.

This chapter is organized into the following sections:

- Overview of the approach used to construct the CDA, a statistically adjusted engineering model.
- Derivation of electric and natural gas engineering estimates, the regression terms in the CDA.
- Specification of the CDA model.
- Model results.

Statistically Adjusted Engineering Analysis

Household energy consumption was decomposed into the demand from various end uses using a regression-based SAE model. Engineering estimates of UECs (engineering UECs) were used as initial point estimates for each end use, such that all end uses in the household had an engineering UEC. The sum of the engineering UECs provided an initial estimate of total consumption for the household. The engineering UECs for each household served as the independent variables in a regression equation, where the dependent variable was the actual energy consumption for each household.

Equation 5-1 provides the general form of the SAE model used to estimate UECs. Separate models were developed for estimating consumption for electricity and natural gas end uses. Household energy consumption was equal to the sum of engineering UECs for all energy-consuming end uses multiplied by scalar adjustment factors (β_i) for each end use, plus residual unexplained error.

Equation 5-1: General Form of SAE Model

$$HHUC_i = \sum_{j=1} (\beta_j * ENG_{ij}) + \varepsilon_{ij}$$

7 *California Statewide Residential Appliance Saturation Study Update to Air Conditioning Unit Energy Consumption Estimates Using 2004 Billing Data*. June 2006. CEC-400-2006-009.

where

$HHUEC_i$ = Energy consumption for household i

ENG_{ji} = Engineering UEC of electricity use for end-use j, for household i.

β_j = Estimated scalar adjustment parameter to the initial UEC for end-use *j*

ε_i = Error term

The scalar adjustments β_j were statistical adjustments made to each engineering UEC. These scalar adjustments were determined as coefficients from the linear regression. A scalar adjustment of one indicated that the engineering UEC provided an exact measure of the amount of energy used by a given end use. A scalar adjustment greater than one indicated the engineering UEC understated the actual consumption derived from the respective end use, i.e., the initial estimate needed to be increased. Conversely, a scalar adjustment of less than one indicated that the engineering UEC overstated the actual consumption derived from the respective end use, i.e., the engineering UEC needed to be decreased. A negative coefficient implied that an end use actually reduced energy consumption.

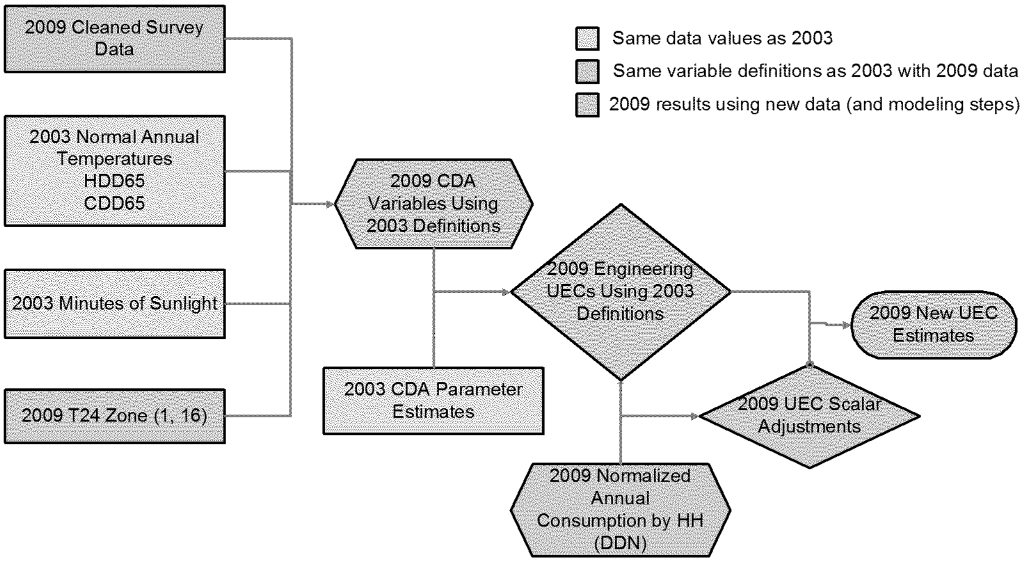
The SAE model for RASS was developed using the process illustrated in Figure 5-1. The data inputs on the left include survey data, normal-year temperatures along with degree-days, and minutes of sunlight. Also, a single indicator variable *T24* identified households in building code zones 1 or 16, which have more restrictive building codes. Survey data provided information on end uses of each fuel, as well as demographic and housing characteristics at the household level. Each of the weather-sensitive end uses was estimated using the same normal temperature and minutes of sunlight series used in the CDA for the 2003 RASS to maintain consistency between the two studies.

The 2009 data inputs were combined to create the CDA variables using the same calculations as for the CDA for the 2003 RASS. The CDA variables consist of linear combinations of appliance and equipment stocks, structural features of the residence, building shell and equipment efficiency factors, weather conditions, and utilization patterns.[8] The 2009 CDA variables were then multiplied by the 2003 CDA parameter estimates and combined to yield initial UEC estimates for each end use. The initial UEC estimates served as the engineering estimates in the SAE model.

---

8 *California Statewide Residential Appliance Saturation Study Final Report*. June 2004. CEC 400-04-009.

Figure 5-1: Overview of SAE Process



Source: *2010 California Residential Appliance Saturation Survey*

The SAE model used normalized annual consumption (NAC) as the dependent variable. The NAC was derived from monthly billing data using the degree-day normalization (DDN) model outlined in Chapter 4. The NAC values for the households were regressed against their respective engineering UECs to provide the scalar adjustments. The scalar adjustments for each end use were multiplied by the corresponding initial engineering estimates to provide adjusted UEC estimates for each end use.

While not shown in the Figure 5-1, the final step of the process was to calibrate the adjusted UECs so that the sum of the final UECs was equal to the observed total NAC. This calibration was done at the sampling-strata level, which included information identifying the electric utility, presence of electric heat, and home type.

The SAE model was estimated using only full-year residents, but the final 2009 UEC estimates contained both full-year and partial-year residents. The final new 2009 UEC estimates were calibrated to average annual consumption by sampling strata from the combined series of full-year and partial-year residents.

Derivation of End-Use Engineering Estimates

CDA equations from the 2003 RASS were used to develop the engineering estimates, which allowed the new 2009 UEC estimates to be directly comparable to the 2003 estimates. Derivation of the UEC estimates is contained in the 2003 report.[1] This section presents the resulting UEC formulas that identify the source of the engineering estimates for the SAE model.

SB_GT&S_0771270

Electric End-Use Engineering Estimates

Engineering estimates were derived for each of the electric end uses listed below.

- Primary space heating
- Secondary space heating
- Central air conditioning
- Room air conditioning
- Evaporative coolers
- Water heating
- Primary refrigerators
- Secondary refrigerators
- Freezers
- Ranges and ovens
- Microwave ovens
- Dishwashers
- Clothes washers
- Dryers
- Outdoor lighting
- Televisions
- Home offices
- Personal computers
- Swimming pool pumps
- Spa pumps
- Spa heat
- Well pumps
- Forced air fans
- Miscellaneous

Each engineering UEC was the sum of one or more cross-product terms, times the corresponding 2003 CDA parameter estimates. The cross-product terms were products of binary variables that indicated the presence of each end use and basic quantitative variables, such as surface area, heating or cooling degree-days, temperature setting, number of units, usage information, or preset engineering parameters. Additional continuous variables differentiated UECs for households according to income level and number of residents in households. Indicator variables were also used to provide separate UEC estimates by residence type, the presence of dual-paned windows, and seasonal effects. While cross-product terms for some of the simpler engineering equations reduced to just one variable, a number of them consisted of multiple cross-product terms with a number of variables in each term. Because only a small number of the variables were relevant to a given household, the sum of these terms was essentially a basic multiplicative formula, with varying adjustments applied depending on the relevant variables.

For example, the general form of the formula for primary space heating is defined in Equation 5-2. The product (1/EFFH)*(DHEAT) consists of the binary variable *DHEAT* that indicated whether electric heat was present in the household, and the term *(1/EFFH)* was an efficiency factor that separates households with conventional electric heat from those with a heat pump. The term *A* denotes a vector overall adjustment, depending on factors such as dwelling type, new construction, and thermostat settings. Finally, the term *B* represents the surface area of the home, the number of heating degree-days, or to the product of the two.

Equation 5-2: General Form of Primary Space Heating

$$\text{Space Heating} = (1/EFFH) * (DHEAT) * A * B$$

Each of the electric end-use engineering UECs is presented below along with a brief description of some of the variables specific to each UEC. The following variables were used in multiple engineering UECs. Detailed descriptions for each of the individual variables used are in Chapter 4.

- HDD65 – Normal heating degree-days from 2003 RASS with a base of 65 degrees
- CDD65 – Normal cooling degree-days from 2003 RASS with a base of 65 degrees
- AREA – Surface area of the residence
- DPWIN – Indicator variable for dual-paned windows
- MF – Indicator variable for multifamily residence
- INC – Continuous variable for household income
- WINTER – Constant adjustment for proportion of winter months
- SUMMER – Constant adjustment for proportion of summer months
- (LOG NUMI + 1) – The number of people in the household, entering the equation in the form (1+log of the number)
- T24 – Indicator variable for household located in building code zone 1 or 16

Space Heating

Primary electric space-heating engineering UECs were developed for both conventional electric *(EHT_ENG)* heat and electric heat with heat pump *(EHP_ENG)*. In addition, engineering UECs were also developed for secondary (or auxiliary) space heating systems (*AUXEH_ENG*) Table 5-1 presents the specific equations used to obtain engineering estimates for each of the space heating end uses.

Table 5-1: Space Heating Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
Primary heating (EHT_ENG and EHP_ENG)	eht_sq_h	3.30E-05	(1/EFFH)*DHEAT*HDD65*AREA
	eht_sq_h_dwp	-8.39E-05	(1/EFFH)*DEHEAT*HDD65*AREA*DPWIN
	eht_sq_h_mf	-0.00112	(1/EFFH)*DEHEAT*HDD65*AREA*MF
	eht_sq_h_inc	-2.90E-10	(1/EFFH)*DEHEAT*HDD65*AREA*INC
	eht_sq_h_inc_dwp	1.77E-10	(1/EFFH)*DEHEAT*HDD65*AREA*INC*DPWIN
	eht_sq_h_inc_mf	2.01E-11	(1/EFFH)*DEHEAT*HDD65*AREA*INC*MF
	eht_sq_h_rm	-3.42E-05	(1/EFFH)*DEHEAT*HDD65*AREA*ROOM
	eht_sq_h_rm_dwp	2.35E-05	(1/EFFH)*DEHEAT*HDD65*AREA*ROOM*DPWIN
	eht_sq_h_rm_mf	1.54E-04	(1/EFFH)*DEHEAT*HDD65*AREA*ROOM*MF
	eht_sq_h_sbk	-7.48E-06	(1/EFFH)*DEHEAT*HDD65*AREA*SETBK
	eht_sq_h_sbk_dwp	-1.52E-05	(1/EFFH)*DEHEAT*HDD65*AREA*SETBK*DPWIN
	eht_sq_h_sbk_mf	5.88E-05	(1/EFFH)*DEHEAT*HDD65*AREA*SETBK*MF
	eht_sq_h_set	3.50E-06	(1/EFFH)*DEHEAT*HDD65*AREA*HTTSET
	eht_sq_h_set_dwp	-1.64E-07	(1/EFFH)*DEHEAT*HDD65*AREA*HTTSET*DPWIN
	eht_sq_h_set_mf	1.86E-05	(1/EFFH)*DEHEAT*HDD65*ARE*HTTSET*MF
	eht_sq_h_nonebu	4.83E-05	(1/EFFH)*DEHEAT*HDD65*ARE*HTTSET*MF*NONELEEK
	eht_sq_winter	0.18559	(1/EFFH)*DEHEAT*AREA*WINTER
	eht_sq_winter_minsun	-2.55E-04	(1/EFFH)*DEHEAT*AREA*WINTER*MINSOFLIGHT
	eht_sq_T24_h	-4.06E-05	(1/EFFH)*DEHEAT*AREA*HDD65*T24
	eth_H_Seasonal	-0.15854	(1/EFFH)*DEHEAT*HDD65*SEASONAL
Secondary Heating (AUXHT_ENG)	eht_aux_h	0.01261	DEAUXHT*HDD65
	eht_aux_sq_h	3.40E-05	DEAUXHT*HDD65*AREA
	eht_aux_sq_h_mf	-1.02E-05	DEAUXHT*HDD65*AREA*MF
	eht_aux_sq_h_freq	1.78E-06	DEAUXHT*HDD65*AREA*ADDFREQ

Source: *2010 California Residential Appliance Saturation Survey*

The primary heating engineering estimates identified households with conventional heating systems and those with systems with a heat pump by using an efficiency factor (1/EFFH). Households with a heat pump were assigned a value of 0.5 for the efficiency factor, thereby reducing the space heating engineering UEC. This was done to reflect greater efficiency derived from heat pump systems than conventional electric heat. Households with conventional heating systems were assigned a value of 1.0 for the efficiency factor, thereby maintaining the higher engineering UEC.

For both primary and auxiliary heating systems, the main driver of the engineering estimates was climate zone-specific heating degree-days with a base temperature of 65 degrees (HDD65). The normal HDD65 series from the 2003 RASS was used to develop engineering estimates to maintain consistency between the two studies. Primary heating system estimates contained additional terms used to adjust the impact of HDD65 on heating usage, depending on the minutes of sunlight, winter months, and whether the residence was a seasonal residence. Additional terms included in the primary heating system estimates allowed for variation in the thermostat setting, building shell, dwelling type, and household income level. Variation in auxiliary heating system engineering estimates was limited to differences in surface area of the residence, dwelling type, and thermostat setting (*ADDFREQ*).

Space Cooling

Space cooling engineering estimates were developed for central air conditioning (CAC_ENG), room air conditioners (RAC_ENG), and evaporative (swamp) coolers (SWAMP_ENG). Table 5-2 presents the specific equations used to obtain engineering estimates for each of these end uses. Many of the terms used in space heating engineering estimates were also used for cooling, but in place of HDD65, the variable CDD65 was used to represent normal cooling degree-days, with a base temperature of 65 degrees. In addition, the central air conditioning equation included a term for new homes. As with the normal HDD65 series, the CDD65 series from the 2003 RASS was used for this study.

Table 5-2: Space Cooling Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
Central Air ConditioningCAC_ENG	cac_sq_c	0.00149	DCAC*CDD65*AREA
	cac_sq_c_new	4.85E-05	DCAC*CDD65*AREA*NEWHOME
	cac_sq_c_dwp	-1.20E-04	DCAC*CDD65*AREA*DPWIN
	cac_sq_c_mf	1.05E-03	DCAC*CDD65*AREA*MF
	cac_sq_c_inc	9.42E-11	DCAC*CDD65*AREA*INC
	cac_sq_c_inc_new	-1.68E-10	DCAC*CDD65*AREA*INC*NEWHOME
	cac_sq_c_inc_dwp	1.25E-10	DCAC*CDD65*AREA*INC*DPWIN
	cac_sq_c_inc_mf	-2.11E-09	DCAC*CDD65*AREA*INC*MF
	cac_sq_c_tset	-1.52E-05	DCAC*CDD65*AREA*TSETC
	cac_sq_c_tset_new	-2.14E-07	DCAC*CDD65*AREA*TESTC*NEWHOME
	cac_sq_c_tset_dwp	9.03E-07	DCAC*CDD65*AREA*TSETC*DPWIN
	cac_sq_c_tset_mf	-1.01E-05	DCAC*CDD65*AREA*TSETC*MF
	cac_sq_minsun_sum	1.00E-04	DCAC*AREA*MINSOFLIGHT*SUMMER
	cac_sq_evp_sum	0.01272	DCAC*AREA*DSWAMP*SUMMER
	cac_c_evp_sq	-1.69E-04	DCAC*CDD65*DSWAMP*AREA
	cac_sq_sum	-0.07495	DCAC*AREA*SUMMER
Room Air Conditioning (RAC_ENG	rac_sq_c	5.15E-05	DRAC*CDD65*AREA
	rac_sq_c_dwp	-1.87E-05	DRAC*CDD65*AREA*DPWIN
	rac_sq_c_mf	1.13E-05	DRAC*CDD65*AREA*MF
	rac_sq_c_inc	-5.83E-10	DRAC*CDD65*AREA*INC
	rac_sq_c_tsetu	1.81E-05	DRAC*CDD65*AREA*TSETUSE
	rac_sq_c_rcnt	1.60E-05	DRAC*CDD65*AREA*RACCNT
	rac_c_evp_sq	-8.93E-05	DRAC*CDD65*DSWAMP*AREA
Evaporative coolers (SWAMP_ENG)	swamp_sq_c	6.35E-05	DSWAMP*AREA*CDD65
	swamp_c	0.19156	DSWAMP*CDD65

Source: *2010 California Residential Appliance Saturation Survey*

Table 5-2 shows two sets of thermostat settings used for space cooling. The variable *TSETC* referred to the average cooling temperature for central air conditioning, while *TSETUSE* was the frequency in which room air conditioners were used.

Water Heating

Engineering estimates were derived for both conventional electric water heating (WHT_ENG) and solar water heating with electric backup (WHTS_ENG). Table 5-3 presents the equation for the water heating engineering estimates, which were distinguished by the presence (or absence) of a tank with solar. The primary driver of the water heating UEC was the number of people in the household, shown by entering the equation in the form (1+log of the number). The number of people in the household also factored into dishwasher usage, clothes washer usage, and the number of showers taken per day.

The equation also included a measure of the average temperature difference from month to month. Because the 2009 RASS used degree-day normalized **annual** consumption data as opposed to **monthly** consumption data, monthly temperature differences were not present in the dataset. Therefore, the average monthly temperature difference by climate zone from the 2003 RASS was used for *WHTEMP_DIFF*. The FACTAWH term from the 2003 CDA was used to adjust for seasonal variation; this variable was equal to a constant for the 2009 CDA because the 2009 CDA was based on annual consumption data.

Table 5-3: Water Heating Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
WHT_ENG and WHTS_ENG	ewh_dwash	28.89343	DEWH*FACTAWH*DWASHU
	ewh_cwash	9.98225	DEWH*FACTAWH*CWASHU
	ewh_shw	18.4293	DEWH*FACTAWH*WHTSHWRS
	ewh_solar	-127.56103	DEWH*FACTAWH*DWHSOLAR
	ewh_add	15.96034	DEWH*ADDWHEL*FACTAWH
	ewh_num	42.08176	DEWH*FACTAWH*Log(NUMI+1)
	ewh_num_mf	-73.10609	DEWH*FACTAWH*Log(NUMI+1)*MF
	ewh_difftemp1	0.03581	DEWH*FACTAWH*WHTEMP_DIFF
	ewh	73.0256	DEWH*FACTAWH

Source: *2010 California Residential Appliance Saturation Survey*

Refrigerators and Freezers

Refrigerator and freezer engineering estimates were primarily based on the Association of Home Appliance Manufacturers (AHAM) engineering estimates as discussed in Chapter 4. Table 5-4 presents the equations for refrigerators and freezers. The second refrigerator estimate also contained terms for variation during the summer months and for multifamily residences.

Table 5-4: Refrigerator and Freezer Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
First Refrigerator (RF1_ENG)	ref1_use	0.0833	DRF1*REFUSAGE1
Second Refrigerator (RF2_ENG)	ref2_use	0.1366	DRF2*REFUSAGE2
	ref2_use_sum	-0.00404	DRF2*SUMMER*REFUSAGE2
	ref2_use_mf	-0.053	DRF2*REFUSAGE2*MF
Freezer (FZ_ENG)	fz_use	0.12464	DFRZR*FZUSAGE

Source: *2010 California Residential Appliance Saturation Survey*

Kitchen

Kitchen appliance engineering estimates were developed for ranges and conventional ovens (RNG_ENG), microwaves (MW_ENG), and dishwashers (DWH_ENG). Table 5-5 presents the equations for ranges, microwaves, and dishwashers. The primary driver for each of these appliances was the number of people in the household, shown by entering the equation in the form (1+log of the number). Engineering estimates for ranges and ovens were allowed to also vary by income and by the presence of a microwave oven. The 2003 CDA used the FACTAMI term to adjust for seasonal variation. For the 2009 CDA, FACTAMI was equal to a constant because the 2009 CDA was based on annual consumption data.

Table 5-5: Kitchen Appliance Engineering Estimates

	CrossProduct Variable	Parameter from 2003 RASS	Equation
Range oven (RNG_ENG)	ecook_num	37.1557	DERNGOV*Log(NUMI+1)
	ecook_num_inc	5.20E-05	DERNGOV*Log(NUMI+1)*INC
	ecook_num_micor	-5.78601	DERNGOV*Log(NUMI+1)*MICRO
	erngov	-22.0967	DERNGOV
Micro-wave (MW_ENG)	micwv	8.33	DMWV*FACTAMI*Log(NUMI+1)
Dishwasher (DWH_ENG)	edwash_num	9.89775	DDW*Log(NUMI+1)*FACTADW
	edw	-6.41515	DDW*FACTADW

Source: *2010 California Residential Appliance Saturation Survey*

Laundry

Laundry included engineering estimates for clothes washers (CWS_ENG) and clothes dryers (EDY_ENG). Equations for these appliances appear in Table 5-6. Similar to kitchen appliances, the primary driver for clothes washers and dryers was the number of people in the household, entering the equation in the form (1+log of the number). The clothes dryer estimate also included a term for the number of loads per day. The FACTACW term was used in the 2003 CDA to adjust for seasonal variation; this variable was equal to a constant for the 2009 CDA because the 2009 CDA was based on annual consumption data.

Table 5-6: Laundry Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
Clothes washer (CWS_ENG)	ecwash_num	37.09798	DCW*FACTACW*Log(NUMI+1)
	ecw	-40.09798	DCW*FACTACW
Clothes Dryer (EDY_ENG)	edry_use	16.78199	DEDRY*FACTADR*EDRYU
	edry_num	5.5022	DEDRY*FACTADR*Log(NUMI+1)
	edry	-27.02423	DEDRY*FACTADR

Source: *2010 California Residential Appliance Saturation Survey*

Outdoor Lighting

The engineering estimate for outdoor lighting (OLT_ENG) was derived using the equation in Table 5-7. The binary variable DOLT indicated whether outdoor lighting was present, while the variable OLTFIX provided a count of the number of outdoor lighting fixtures. The formula allowed for differentiation based on the number of outdoor fixtures that used CFLs and high-intensity discharge (HID) bulbs, or fixtures on sensors and timers. The variable HRDK indicated the number of hours of darkness in the climate zone.

Due to multicollinearity problems, the 2003 RASS and this study did not estimate a separate indoor lighting UEC. Indoor lighting was assumed to be part of the *Miscellaneous UEC*.

Table 5-7: Outdoor Lighting Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
OLT_ENG	olit_cfl	-5.65594	DOLT*OLTFIX*ONOCFL
	olit_hid	5.26879	DOLT*OLTFIX*OPROPHID
	olit_sen	-4.17967	DOLT*OLTFIX*OPROPSENS
	olit_tim	11.10408	DOLT*OLTFIX*OPROPTIM
	olit_hrdk	2.11248	DOLT*OLTFIX*HRDK
	olt	-20.00278	DOLT*OLTFIX

Source: *2010 California Residential Appliance Saturation Survey*

Home Electronics and Office

While separate engineering estimates were derived in the 2009 RASS for televisions (CTV_ENG), home offices (OFF_ENG), and personal computers (PCS_ENG), they are presented together in Table 5-8. The equations for each of these engineering estimates contained terms to include the number of hours of use. In addition, the televisions* estimate assumed a value to differentiate the energy consumption per hour between large- and small-screen televisions. Personal computers included a term for the total number of desktop and laptop computers in the household.

Table 5-8: Home Electronics and Office Equipment Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
TV's (CTV_ENG)	tvhrs_kw	36.48776	DTV*TVKW*TVHRS
	tv_kw	99.84392	DTV*TVKW
Home office (OFF_ENG)	ehmoffuse	0.80713	DHMOFF*HMOFFHRS
	hmoff	-0.712	DHMOFF
Personal computers (PCS_ENG)	pc_num	16.48716	DPC*PCNUM
	pc_num_hrs	1.68823	DPC*PCNUM*PCHRS1
	epc	6.52058	DPC

Source: *2010 California Residential Appliance Saturation Survey*

Swimming Pool and Spa

Engineering estimates were developed for swimming-pool filter pumps (PMP_ENG), spa filter (SPA_ENG) pumps, and spa heaters SPH_ENG). Each of these engineering estimates was a function of the frequency of use, as shown in Table 5-9. Swimming-pool filter use was indicated by PLFILT, while spa filter and spa heat use were indicated by SPAFREQ and SPAHTFRQ, respectively. Additional terms were added to the spa heat estimate to account for a cover or a combined electric and solar spa heating system.

Table 5-9: Swimming Pool and Spa Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
Pool filter pump (PMP_ENG)	plpmp_flt	-17.9017	DPLPMP*PLFILT
	plpmp_flt_sz	0.00116	DPLPMP*PLFILT*PLSIZE
	plpmp	177.43949	DPLPMP
Spa filter pump (SPA_ENG)	spa_pmp	1.8575	DSPA*SPAFREQ
	spa_pmp_sz	0.6434	DSPA*SPAFREQ*SPASIZE
Spa heat (SPH_ENG)	espa_ht_freq	4.11848	DEHTSPA*SPAEHTFREQ
	espa_ht_freq_sz	-0.19491	DEHTSPA*SPAEHTFREQ*SPASIZE
	espa_ht_sz_cov	7.22828	DEHTSPA*SPASIZE*SPCOV
	espa_ht_solar	6.29138	DEHTSPA*SPASOLAR

Source: *2010 California Residential Appliance Saturation Survey*

Well Pump

Well pumps (WPM_ENG) are used in areas that do not have municipal water. The well-pump engineering estimate was based entirely on the number of people in the household, entering the equation in the form (1+log of the number), as shown in Table 5-10.

Table 5-10: Well-Pump Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
WPM_ENG	wellpuse	55.41209	DWELLP*Log(NUMI+1)
	wellp	0.64884	DWELLP

Source: *2010 California Residential Appliance Saturation Survey*

Forced Air Fan

The CDA for the 2003 RASS included an estimate for forced-air furnace fans (VENT1_ENG). Table 5-11 shows that this engineering estimate was based solely on HDD65 and the surface area of the home.

Table 5-11: Forced-Air Fan

	Cross Product Variable	Parameter from 2003 RASS	Equation
VENT1_ENG	fafan_sq_h	2.30E-05	DFFAN*HDD65*AREA

Source: *2010 California Residential Appliance Saturation Survey*

Miscellaneous Uses

The engineering estimate for the miscellaneous UEC (MISC_ENG) was estimated using the equation presented in Table 5-12. It accounted for all energy consumption not captured by the other UECs. The terms used in this engineering estimate include a combination of demographic, structural, and seasonal variables. In addition, parameters for attic or ceiling fans were also included to avoid collinearity the cooling terms.

Table 5-12: Miscellaneous Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
MISC_ENG	miss_inc	3.09E-04	INC
	miss_sq	0.04769	SQFT
	miss_numi	43.11824	Log(NUMI+1)
	miss_newh	-42.01492	NEWHOME
	miss_mf	-8.54592	MF
	miss_seasonal	-142.36973	SEASONAL
	miss_ceil	19.19172	DCEILF
	fat_c	0.35164	DATTFAN*CDD65
	fat_sq_c	-7.05E-05	DATTFAN*CDD65*AREA
	miss_epl_ht	88.18653	EPLHT

Source: *2010 California Residential Appliance Saturation Survey*

Natural Gas End-Use Engineering Estimates

Engineering estimates were derived for each of the natural gas end uses listed below.[9] While a general description of the formulas used to create estimates for each end use is presented below, derivation of these formulas can be found in the 2003 RASS report.[1]

- Primary space heating
- Secondary space heating
- Water heating
- Ranges and ovens
- Clothes dryers
- Swimming pools and spas
- Miscellaneous

Space Heating

Table 5-13 presents the equations for primary (GHT_ENG) and secondary (auxiliary) (GAUXHT_ENG) natural gas space heating. The terms used to estimate primary natural gas heating were similar to those used for electric heating estimates but also included terms for system age and whether the residence was a new home or mobile home.

9 Note the 2003 RASS report misprinted parameters for some of the terms in the 2003 natural gas CDA. Estimates presented reflect the correct CDA results.

Table 5-13: Space Heating Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
Primary heating (GHT_ENG)	ght_sq_T24_winter	0.000238	DGHEAT*AREA*WINTER*T24
	ght_sq_T24_h	-1.60E-05	DGHEAT* HDD65* AREA *T24
	ght_sq_h	-2.68E-07	DGHEAT*HDD65* AREA
	ght_sq_h_new	-6.70E-06	DGHEAT*HDD65* AREA *NEWHOME
	ght_sq_h_age	-1.90E-06	DGHEAT*HDD65* AREA *GHTAGE
	ght_sq_h_dwp	-2.50E-06	DGHEAT*HDD65* AREA *DPWIN
	ght_sq_h_mf	-4.00E-05	DGHEAT*HDD65* AREA *MF
	ght_sq_h_inc	4.73E-11	DGHEAT*HDD65* AREA *INC
	ght_sq_h_inc_new	6.42E-12	DGHEAT*HDD65* AREA *INC*NEWHOME
	ght_sq_h_inc_age	-6.31E-13	DGHEAT*HDD65* AREA *INC*GHTAGE
	ght_sq_h_inc_dwp	-1.97E-11	DGHEAT*HDD65* AREA *INC*DPWIN
	ght_sq_h_inc_mf	-1.11E-11	DGHEAT*HDD65* AREA *INC*MF
	ght_sq_h_rm	2.26E-06	DGHEAT*HDD65* AREA *GROOM
	ght_sq_h_rm_age	-3.13E-07	DGHEAT*HDD65* AREA *GROOM*GHTAGE
	ght_sq_h_rm_dwp	4.56E-06	DGHEAT*HDD65* AREA *GROOM*DPWIN
	ght_sq_h_rm_mf	2.27E-06	DGHEAT*HDD65* AREA *GROOM*MF
	ght_sq_h_sbk	-5.18E-07	DGHEAT*HDD65* AREA *SETBK
	ght_sq_h_sbk_age	-1.32E-07	DGHEAT*HDD65* AREA *SETBK*GHTAGE
	ght_sq_h_sbk_dwp	1.73E-06	DGHEAT*HDD65* AREA *SETBK*DPWIN
	ght_sq_h_sbk_mf	4.95E-06	DGHEAT*HDD65* AREA *SETBK*MF
	ght_sq_h_set	5.36E-07	DGHEAT*HDD65* AREA *HTTSET
	ght_sq_h_set_age	3.04E-08	DGHEAT*HDD65* AREA *HTTSET*GHTAGE
	ght_sq_h_set_dwp	-6.13E-08	DGHEAT*HDD65* AREA *HTTSET*DPWIN
	ght_sq_h_set_mf	5.96E-07	DGHEAT*HDD65* AREA *HTTSET*MF
	ght_sq_h_nonebu	-1.70E-06	DGHEAT*HDD65* AREA *NONGBU
	ght_sq_winter	0.01694	DGHEAT*AREA *WINTER
	ght_sq_winter_minsun	-2.30E-05	DGHEAT* AREA *WINTER*MINSOFLIGHT
	ght_h_age	-0.00847	DGHEAT*HDD65*GHTAGE
	ght_h_new	0.00104	DGHEAT*HDD65*NEWHOME
	gth_h_seasonal	-0.00771	DGHEAT*HDD65*SEASONAL
	ght_sq_h_mh	5.23E-06	DGHEAT*HDD65*AREA *MH
	ght_sq_h_inc_mh	-4.42E-11	DGHEAT*HDD65*AREA*INC*MH
Secondary Heating (GAUXHT_ENG)	ght_aux_h	0.022054	DNGAUXHT*HDD65
	ght_aux_sq_h	0.000003812	DNGAUXHT*HDD65* AREA
	ght_aux_sq_h_mf	-0.000001903	DNGAUXHT*HDD65* AREA *MF

Source: *2010 California Residential Appliance Saturation Survey*

Water Heating

Table 5-14 presents the equation used to produce engineering estimates of natural gas water heating. The primary differences between the electric and natural gas estimates were the natural gas equation terms for seasonal variation and new homes, while the electric estimate contained terms for multifamily households and whether an electric water heater was added in the past year.

Table 5-14: Water Heating Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
GWHT_ENG and GWHTS_ENG	gwh_num	-1.17111	DGWH*FACTAWH* Log(NUMI+1)
	gwh_dwash	0.65463	DGWH*FACTAWH*DWASHU
	gwh_cwash	0.45847	DGWH*FACTAWH*CWASHU
	gwh_solar	-2.67182	DGWH*FACTAWH*DWHGSOLAR
	gwh_num_new	-3.13922	DGWH*FACTAWH*LOG(NUMI+1)*NEWHOME
	gwh_num_seasonal	-9.0196	DGWH*FACTAWH* Log(NUMI+1)*SEASONAL
	gwh	1.40E+01	DGWH*FACTAWH
	gwh_difftemp1	0.009662794	DGWH*FACTAWH*WHTEMP_DIFF
	gwh_shw	0.21075	DGWH*FACTAWH*TOTAL_SHTSHWRS

Source: *2010 California Residential Appliance Saturation Survey*

Kitchen

Natural gas kitchen appliances were limited to ranges and ovens (GRNG_ENG). Table 5-15 shows the equations used to create engineering estimates for these appliances. The same terms were used as for electric ranges and ovens.

Table 5-15: Range or Oven Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
GRNG_ENG	gcook_num	6.31481	DGRNGOV* Log(NUMI+1)
	gcook_num_inc	-3.11E-06	DGRNGOV* Log(NUMI+1)*INC
	gcook_num_micor	-1.24E+00	DGRNGOV* Log(NUMI+1)*MICRO
	dgrngov	-3.18E+00	DGRNGOV

Source: *2010 California Residential Appliance Saturation Survey*

Laundry

Engineering estimates for natural gas clothes dryers were estimated using the equation shown in Table 5-16. The terms were the same as the ones used for electric clothes dryers.

Table 5-16: Clothes Dryer Engineering Estimates

	Cross Product Variable	Parameter from 2003 RASS	Equation
GDRY_ENG	gdry_use	0.6391	DGDRY*FACTADR*GDRYU
	gdry_num	0.50575	DGDRY*FACTADR* Log(NUMI+1)
	gdry	-1.53717	DGDRY*FACTADR

Source: *2010 California Residential Appliance Saturation Survey*

Swimming Pool and Spa

The equations used to produce engineering estimates for natural gas heat for pools and spas are presented in Table 5-17. Both sets of equations were functions of size and frequency of use, as well as an adjustment for whether the pool or spa had a cover.

Table 5-17: Swimming Pool and Spa Engineering Estimates

	CrossProduct Variable	Parameter from 2003 RASS	Equation
GPHT_ENG	gpl_ht	-1.30781	DGPLHT
	gpl_ht_freq	2.76838	DGPLHT*GPLHTFREQ
	gpl_ht_sz	0.00046	DGPLHT*PLSIZE
	gpl_ht_sz_cov	0.000234	DGPLHT*PLSIZE*DPLCOV
Spa Heat (GSPA_ENG)	gspa_ht	3.5606	DGHTSPA
	gspa_ht_freq	0.81287	DGHTSPA*SPAGHTFREQ
	gspa_ht_freq_sz	0.00161	DGHTSPA*SPAGHTFREQ*SPASIZE
	gspa_ht_sz_cov	-0.12805	DGHTSPA*SPASIZE*SPCOV
	gspa_ht_solar	1.64078	DGHTSPA*SPAGSOLAR

Source: *2010 California Residential Appliance Saturation Survey*

Miscellaneous

Table 5-18 shows the engineering estimate for the natural gas miscellaneous UEC (GMISC_ENG) that contained terms for medical equipment and natural gas barbecues.

Table 5-18: Miscellaneous Engineering Estimates

	CrossProduct Variable	Parameter from 2003 RASS	Equation
GMISC_ENG	miss_gmedical	2.70E+01	DGMED
	miss_gbbq	2.22319	DGBBQ

Source: *2010 California Residential Appliance Saturation Survey*

Specification of CDA Models

The engineering estimates presented above were used to construct separate electric and natural gas SAE models. The basic model consisted of linear combinations of the respective electric and natural gas engineering estimates for each household. The intercept term was excluded from each model, thereby constraining household consumption to equal the sum of the individual engineering estimates plus residual error.

Due to collinearity among end-use terms, a number of end uses were combined in both the electric and natural gas models. For combined terms, a single scalar adjustment was estimated and applied to each of the individual UECs.

The derivation of the electric and natural gas SAE models are presented separately below.

Derivation of the Electric SAE

This section presents the specification of the electric SAE model. While many of the engineering estimates entered the model directly or as a binary variable multiplied by the engineering estimate, some end uses required additional manipulation. Interaction terms were used to adjust primary electric heat and central air conditioning estimates for households without heating and cooling DDN terms. In addition, a number of engineering estimates were combined due to multicollinearity. The treatment of each engineering estimate in the SAE model is discussed below.

Space Heating

The SAE model estimated scalar adjustments for both primary and secondary electric heating systems. Primary electric heat entered the SAE model as two terms, with separate scalar adjustments for each term. The first term consisted of the combined engineering estimates for conventional electric heat and electric heat with a heat pump. Each of these engineering estimates was interacted with a binary variable (*DEHEAT*), identifying whether an electric heating system was present in the household. Because households with electric heat were not allowed to have both primary conventional electric heat and electric heat with a heat pump, the combined term resulted in the engineering estimate for the appropriate system for each household. A single scalar adjustment was estimated for the combined electric heat and heat pump term *EHT_ENG_NEW* as shown in Equation 5-3.

Equation 5-3: Electric Space Heating

$$EHT\_ENG\_NEW = DEHEAT(EHT\_ENG + EHP\_ENG)$$

A second primary heating term was used to account for households that both the survey responses and the electric utility indicated that they had primary electric heat, but no significant relationship between consumption and HDD was found by the DDN model. These households were assumed to have electric heat but seldom used it and were identified in the model using a second term depicted in Equation 5-4. The equation shows the electric heat engineering estimate (*EHT_ENG*) interacted with the binary variable (*NG_NP_UE*). As defined in Chapter 4, *NG_NP_UE* identified households that did not have natural gas service or a significant heating term from the electric DDN model, but the utility identified them as having a heating allowance. Separate scalar adjustments were estimated for *EHT_ENG_NEW* and *EHT_NoDDN*.

Equation 5-4: Electric Space Heating With No Heating Term From the DDN

$$EHT\_NoDDN = DEHEAT * EHT\_ENG * NG\_NP\_UE$$

A separate term was added to the SAE model for secondary (or auxiliary) space heating. As shown in Equation 5-5, this term was set as equal to the engineering estimate (*AUXHT_ENG*) times the binary variable (*DEAUXHT*) indicating the presence of electric auxiliary heat.

Equation 5-5: Electric Auxiliary Space Heating

$$AUXHT\_ENG\_NEW = DEAUXHT * AUXHT\_ENG$$

Space Cooling

Space cooling end uses consisted of central air conditioning, room air conditioning, and evaporative (swamp) coolers. The central air conditioning SAE adjustment consisted of two terms, seen in Equations 5-6 and 5-7. The first term was simply an interaction of engineering estimates for central air conditioning with an indicator variable for people who had central air conditioning. The second term added an additional interaction effect for households that reported having central air, but did not have a significant cooling term from the DDN model. This term accounted for people who have central air, but do not use it frequently. Separate scalar adjustments were estimated for CAC_*ENG_NEW* and CAC_*NoDDN*

Equation 5-6: Central Air Conditioning

$$CAC\_ENG\_NEW = CAC\_ENG * DCAC$$

Equation 5-7: Central Air Conditioning with No Cooling Term from the DDN

$$CAC\_NoDDN = CAC\_ENG * DCAC * NoDDN$$

In addition, separate terms were also added for room air conditioning and evaporative cooler. As seen in Equations 5-8 and 5-9, each of these terms simply consisted of the respective engineering estimate times a binary variable indicating the presence of either a room air conditioner or evaporative cooler.

Equation 5-8: Room Air Conditioning

$$RAC\_ENG\_NEW = RAC\_ENG * DRAC$$

Equation 5-9: Evaporative Cooling

$$SWAMP\_NEW = SWAMP\_ENG * DSWAMP$$

Water Heating

The engineering estimates for stand-alone electric water heating and solar water heating with electric backup were combined, as seen in Equation 5-10.

Equation 5-10: Electric Water Heating

$$WHT\_ENG\_NEW = WHT\_ENG * DEWHT + WHTS\_ENG * DEWHTSOLAR$$

Kitchen Appliances

The range/oven, microwave, and dishwasher engineering estimates were collapsed into a single variable *KITCHEN*. The estimated scalar adjustment for KITCHEN was applied to engineering estimates for each respective appliance.

Equation 5-11: Electric Kitchen Appliances

$$KITCHEN = RNG\_ENG * DERNGOV + MW\_ENG * DMW + DWH\_ENG * DWH$$

Laundry Equipment

A new variable *LAUNDRY* was derived from the sum of clothes washer and electric clothes dryer engineering estimates, as seen in Equation 5-12.

Equation 5-12: Electric Laundry Equipment

$$LAUNDRY = CWS\_ENG * DCWS + EDY\_ENG * DEDRY$$

Spas

Estimates for spa filter pumps (*SPA_ENG*) and electric spa heaters (*SPH_ENG*) were combined into the single variable *SPA*. The new variable *SPA* was the sum of spa filter and spa heating engineering estimates multiplied by their respective binary variables. A single scalar adjustment was estimated for the term *SPA*.

Equation 5-13: Spa Filter Pumps and Electric Spa Heating

$$SPA = SPA\_ENG * DSPA + SPH * DESPAH$$

Miscellaneous and Ventilation

The 2003 RASS estimated separate UECs for forced-air fans and attic/ceiling fans, but forced-air fans were estimated separately. However, the SAE model also combined the forced-air fan (*VENT1_ENG*) with the miscellaneous engineering estimate (*MISC_ENG*), providing a single scalar adjustment for the combined term.

Equation 5-14: Electric Miscellaneous and Ventilation

$$MISC\_ENG\_NEW = MISC\_ENG * DMISC + VENT1\_ENG * DVENT1$$

Other Electric SAE Terms

The remaining electric end uses were included in the SAE model by simply multiplying the engineering estimates by indicator variables that identified the presence of each end use. Each of the terms listed Table 5-19 received a separate scalar adjustment.

Table 5-19: Electric End-Use Terms for SAE Model

Appliance	SAE Term
Outdoor Lights	OLTUSE_NEW = OLTUSEeng * DOLT
Televisions	TVUSE_NEW = TVUSEeng * DTV
Home Office	OFFUSE_NEW = OFFUSEeng * DHMOFF
PersonalComputer	PCS_ENG_NEW= PCS_ENG* DPC
Pool filter pump	PMP_ENG_NEW= PMP_ENG* DPLPMP
Well pump	WPM_ENG_NEW= WPM_ENG* DWELLP
First refrigerator	RF1_ENG_NEW = RF1_ENG * DRF1
Second refrigerator	RF2_ENG_NEW = RF1_ENG * DRF2
Freezer	FZ_ENG_NEW = FZ_ENG * DFZ

Source: *2010 California Residential Appliance Saturation Survey*

Summary of SAE Electric Model

In summary, the final SAE electric model was expressed as the following equation.

Equation 5-15: SAE Electric Model

$$NAC_{kwh} = \beta_1 * EHT\_ENG\_NEW + \beta_2 * EHT\_\mathit{NoDDN} + \beta_3 * AUXHT\_ENG\_NEW +$$

$$\beta_4 * \mathit{CAC\_ENG\_NEW} + \beta_5 * \mathit{CAC\_NoDDN} + \beta_6 * \mathit{RAC\_ENG\_NEW} +$$

$$\beta_7 * \mathit{SWAMP\_NEW} + \beta_8 * \mathit{WHT\_ENG\_NEW} + \beta_9 * \mathit{KITCHEN}$$

$$\beta_{10} * \mathit{LAUNDRY} + \beta_{11} * \mathit{SPA} + \beta_{12} * \mathit{OLTUSE\_NEW} + \beta_{13} * \mathit{TVUSE\_NEW}$$

$$\beta_{14} * \mathit{OFFUSE\_NEW} + \beta_{15} * \mathit{PCS\_ENG\_NEW} + \beta_{16} * \mathit{PMP\_ENG\_NEW}$$

$$\beta_{17} * \mathit{WPM\_ENG\_NEW} + \beta_{18} * \mathit{RF1\_ENG\_NEW} + \beta_{19} * \mathit{RF2\_ENG\_NEW} +$$

$$\beta_{20} * \mathit{FZ\_ENG\_NEW} + \beta_{21} * \mathit{MISC\_ENG\_NEW}$$

Derivation of the Natural Gas SAE

The natural gas SAE was limited to three terms: space heating, water heating, and base load. The derivation of these terms is presented below.

Space Heating

There was a relatively high cross-saturation of natural gas space heating and water heating. Furthermore, because the SAE model was based on annualized consumption data, isolating the correct scalar adjustments for these two terms required additional information to be added to the SAE model. Specifically, the predicted heat-sensitive load from the DDN model was incorporated into the natural gas space heating SAE term as discussed below. This allowed the model to identify the proportion of the overall adjustment that was weather-sensitive and therefore provisionally attributed to space heating. The following process was used to develop the natural gas space heating term.

First, engineering estimates for primary and secondary natural gas space heating were combined, as seen in Equation 5-16. The equation shows engineering estimates for primary and secondary natural gas heating each multiplied by their respective binary variable.

Equation 5-16: Natural Gas Primary and Auxiliary Space Heating

$$\mathit{GHT\_ENG\_NEW = GHT\_ENG * DGHEAT + GAUXHT\_ENG * DNGAUXHT}$$

For households that had a predicted heating load from the DDN model, the heating load portion of their normalized consumption and GHT_ENG_NEW were averaged, as shown in Equation 5-17.

Equation 5-17: Natural Gas Space Heating With Heating Load From DDN

$$\mathit{GHT\_ENG\_NEW1} = (\mathit{GHT\_ENG\_NEW}_w + \text{Predicted heating load from DDN}) / 2$$

SB_GT&S_0771287

For households that did not have a heating load from the DDN model, the value of GHT_ENG_NEW from Equation 5.15 was used as their new engineering estimate for natural gas heat. Therefore, GHT_ENG_NEW1 was set equal to GHT_ENG_NEW for these households.

Water Heating

Similar to electric water heat, the engineering estimates for standalone natural gas water heat with natural gas backup were combined, as seen in Equation 5-18. As presented in Chapter 4, the variables *DGWHT* and *DGWHT*SOLAR were binary variables indicating whether a household had each respective water heater type.

Equation 5-18: Natural Gas Water Heating

$$GWHT\_ENG\_NEW = GWHT\_ENG * DGWHT + GWHTS\_ENG * DGWHTSOLAR$$

Natural Gas Base Load

The natural gas range/oven, clothes dryer, spa heat, pool heat, and miscellaneous engineering estimates were combined to make up the BASE term, as seen in Equation 5-19. This term received a single scalar adjustment in the SAE model.

Equation 5-19: Natural Gas Base Load

$$GBASEUSE_{new} = GRNGOVUSE_{eng} * DGRNGOV + GDRYUSE_{eng} * DGDRY + GSPAHUSE_{eng} * DGSPAH + GPLHUSE_{eng} * DGPLH$$

Summary of SAE Natural Gas Model

In summary, the final SAE natural gas model was expressed as the following equation.

Equation 5-20: SAE Natural Gas Model

$$NAC_{therms} = \beta_1 * GHT\_ENG\_NEW1 + \beta_2 * GWHT\_ENG\_NEW + \beta_3 * GBASE\_ENG\_NEW$$

Estimated Model Results

The electric and natural gas SAE models were estimated using an ordinary least squares method for households with a fitted DDN model.

As seen in Table 5-20, all parameter estimates in the final electric model were significant within 0.05%. The scalar adjustments for most coefficients were within 20 percent of the engineering estimates. Only outdoor lighting and televisions had scalar adjustments that increased initial engineering UEC estimates by more than 20 percent. However, central air conditioning, spa filter pumps, home offices, and well pumps all had adjustments that decreased the initial engineering estimates by more than 20 percent. Also, the coefficient on first refrigerators was restricted to *1* due to collinearity with second refrigerators and freezers.

Table 5-20: Electric SAE Model

Variable	Parameter Estimate	Standard Error	t Value
Electric Heat (Combined)	0.79472	0.05913	13.44
Electric Heat UEC for households with no gas, no DDN model, but a utility heat allowance	-0.49269	0.09442	-5.22
Auxiliary Heat	0.75463	0.38366	1.97
Central Air	0.7153	0.02154	33.2
Central Air for households with no DDN	-0.48567	0.03274	-14.83
Room AC	1.06379	0.16152	6.59
Swamp Cooler	0.86872	0.09461	9.18
Kitchen	0.98652	0.11605	8.5
Spa	0.62431	0.03879	16.1
Laundry	0.98693	0.04906	20.12
Water Heat	0.94346	0.03144	30.01
Outdoor Lights	1.30454	0.06292	20.73
TV	1.21233	0.05614	21.6
Home Office	0.64009	0.28992	2.21
PC's	0.84537	0.03892	21.72
Pool Pump	1.18914	0.02607	45.62
Well Pump	0.62557	0.11746	5.33
Misc	0.99635	0.02602	38.28
Refrigerator 1	1	0	Infinity
Refrigerator 2	1.01043	0.03954	25.55
Freezer	1.06159	0.05483	19.36
RESTRICT	-323,553,462	147,727,543	-2.19

Source: *2010 California Residential Appliance Saturation Survey*

The three parameters in the natural gas model, seen in Table 5-21, were also highly significant. The natural gas heating parameter indicates a 24 percent reduction in the natural gas heating UEC from the initial engineering estimate. However, the adjusted estimate of water heating use was almost exactly equal to the initial engineering estimate. The base consumption parameter was a 20 percent reduction from the initial engineering estimates of all terms in the base load.

Table 5-21: Gas SAE Model

Variable	Parameter Estimate	Standard Error	t Value
Gas Heat	0.76083	0.0088	86.47
Water Heat	0.99817	0.01391	71.78
Base	0.80722	0.01962	41.14

Source: *2010 California Residential Appliance Saturation Survey*

List of Acronyms

CAC	central air conditioning
CDA	conditional demand analysis
CFL	compact fluorescent lamp
DDN	degree-day normalization
Energy Commission	California Energy Commission
F	Fahrenheit
IOU	investor-owned utilities
LADWP	Los Angeles Department of Water and Power
LCD	liquid crystal display
NAC	normalized annual consumption
PC	personal computer
PG&E	Pacific Gas and Electric Company
RAC	room air conditioning
RASS	Residential Appliance Saturation Survey
SAE	statistically adjusted engineering
SAS	statistical analysis system
SCE	Southern California Edison Company
SDG&E	San Diego Gas & Electric Company
SoCal Gas	Southern California Gas Company
UEC	unit energy consumption