1997 Residential Energy Management Services
First Year Load Impact Evaluation
(Home Energy Fitness Program)

Study ID No. 715

Final Report

Prepared by

Andrew A. Goett
AAG & Associates

Prepared for
Southern California Gas Company

March 1999

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report presents the results of the first year impact evaluation of Southern California Gas Company's 1997 Residential Energy Management Services Program, which is known as the Home Energy Fitness Program (HEF Program). The HEF Program promotes the adoption of energy efficient measures and actions by providing informational audits to residential customers. The objective of the impact evaluation was to estimate the natural gas savings that resulted from the 1997 HEF Program audits. This evaluation was performed in compliance with the requirements of the Protocols and Procedures for Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs, March 1998 (M&E Protocols) that govern the procedures that the California investor-owned utilities must use in evaluating their programs.

The impact evaluation estimated the net savings attributable to the 1997 HEF Program by examining the consumption patterns for a representative sample of participants and nonparticipants over a 27-month period spanning the receipt of the audits. Several statistical regression models were estimated on these data, and they produced net annual savings per participant estimates in the range of 16.7 to 28.8 therms per year. The net savings are the reductions in gas consumption from the program, after controlling for "naturally occurring" conservation and background environmental and economic trends.

Based on the performance of the different regression models that were estimated, one was selected as the preferred specification to calculate net program savings. This model produced a net annual savings estimate of 28.8 therms per participant. Given this model, the reliability of the estimates is fairly high. With a statistical confidence of 90%, one can state that the "true" net program savings are at least 17.9 therms and less than 39.7 therms per participant per year.

Table E-1
Net Impact Estimates (Therms)

	Ex Ante Impact[1]	Ex Post Impact	Ex Post # of Audits[2]	Ex Post Program Impact	Realization Rate[3]
Do-It-Yourself Surveys (Audits)	**44**	**28.8**	**22,818**	**657,158**	**65%**

Notes:
1 Ex ante estimate filed in Advice Letter 2526, October 1, 1996
2 Number of surveys processed from January 1, 1997, through December 31, 1997.
3 Load impact estimated by this study (ex post impact) divided by the load impact filed in the first year earnings claim (ex ante impact).

The data preparation and analysis used to derive the estimates, as well as their documentation, adhered to the requirements of the M&E Protocols concerning load impact measurement. The statistical regression analysis is a Load Impact Regression Model (LIRM) as defined in the Protocols. The model is based on accepted empirical techniques for

measuring the impacts of DSM programs, and it produces diagnostics that allow independent assessment of its performance. The LIRM specifications were based on sound behavioral and physical principles, and they included the major variables expected to influence gas consumption. Tests of important statistical issues that could arise in the estimation of model parameters were performed, and appropriate procedures were used to correct any significant problems that were identified.

The data met the requirements of the M&E Protocols concerning participant and nonparticipant sample sizes and coverage. The process of data acquisition and preparation is fully documented in this report and its appendices. All of the regression models that were estimated are presented in the report, with associated confidence statistics and related information.

Chapter 1
INTRODUCTION: IMPACT EVALUATION

FIRST YEAR LOAD IMPACT EVALUATION OBJECTIVE

This report presents the results of the first year load impact evaluation of Southern California Gas (SoCalGas) Company's 1997 Home Energy Fitness (HEF) Program and documents the data and analysis upon which the results are based. The HEF Program promotes the adoption of energy efficient measures and actions by providing informational audits to residential customers.

The objective of the impact evaluation was to estimate the natural gas savings that resulted from the 1997 HEF Program audits. The estimates represent net changes, after accounting for the effects of actions that participants would have taken in the absence of the program and trends in gas consumption.

The data collection and analysis used to obtain the impact results conform to the requirements of the Protocols for the Verification of Demand-Side Management (DSM) Programs (M&E Protocols) governing the procedures used by the California investor-owned utilities to estimate the impacts of their DSM programs. The M&E Protocols specify a series of requirements about data sources, analysis procedures, and reporting for impact evaluations. The key requirements that relate to this report are summarized in Tables 5, 6, 7, and C-11 of the M&E Protocols. Table 5 specifies the measurement methodology, the sample design and size, and the billing data requirements. Table 6 specifies the reporting requirements for the program impacts. Table 7 specifies the documentation requirements for data preparation and analysis. Table C-11 sets out additional requirements for the evaluation of Energy Management Services Programs, of which the HEF Program is an example.

DESCRIPTION OF THE HOME ENERGY FITNESS PROGRAM

SoCalGas' 1997 HEF Program offered residential customers an analysis of their gas consumption and made recommendations as to how they could reduce their usage by adopting cost effective measures and practices. The analysis estimated a residential customer's annual natural gas use by end use based on past consumption data and responses to a questionnaire that the participants completed on their gas using equipment characteristics and operation.

SoCalGas marketed the HEF Program in 1997 primarily by targeted mailings. For its direct mail campaign (200,000), SoCalGas targeted residential, single-family detached homes, condos, or townhouses. The target population did not include mobile home parks (master metered) or apartment dwellers. Apart from the direct mail campaign, customers who called the call center regarding a high bill investigation, or directly requested a survey were sent one if they fell within this criteria.

When pulling the direct mailing list, SoCalGas initially based its selection criteria on at

least 10 years of continuous service. We eliminated customers who had been previously contacted, which resulted in 156,000 addresses. We then lowered the years of continuous service to 5 years to obtain another 44,000 (also eliminated customers who had been previously contacted).

The marketing brochure invited SoCalGas customers to complete an enclosed questionnaire and to receive a report by return mail. The offer letter and typical report are reproduced in Appendix A.

In 1997, SoCalGas sent out approximately 200,000 of the HEF offer letters to its residential customers. Approximately 23,000 customers responded and received a personalized Home Energy Fitness Report. SoCalGas treated these customers as HEF Program participants. This report included a personalized analysis of how much of the customer's annual gas bill was distributed by end use. It also made recommendations regarding how the customer potentially could save natural gas energy and reduce his/her gas bill. In addition, estimates of the cost savings associated with the recommendations were provided to the customer. An example of a typical report and a list of the types of recommendations are shown in Appendix A.

METHODOLOGY OVERVIEW

The estimated savings from the HEF Program were based on a statistical analysis of the gas consumption for a sample of participants and nonparticipants from September 1996 through November 1998. The statistical analysis consisted of the estimation of a series of regression models that relate consumption to variables representing program participation, weather, and gas equipment holdings, in addition to other residence and household characteristics. The specifications used in this analysis are referred to as Load Impact Regression Models (LIRM) in the M&E Protocols.

The models are based on accepted empirical techniques for measuring program impacts, and the coefficients are estimated using techniques that account for the important statistical issues that arise in the estimation of regression model parameters. The estimation method accounted for both autocorrelation and heteroskedasticity, which were found to be present in the data. In addition, other potential statistical problems were investigated. A detailed explanation of the model used to estimate program impacts appears in Chapter 3 of this report.

The data on which the regression models were estimated came from four sources. The first was the 1997 HEF Program tracking database. This identified the participants and the dates on which the audit reports were sent to them. The second data source was SoCalGas' billing system, from which consumption data were extracted for participants and a sample of nonparticipants covering the period from September 1996 through November 1998. The third source was newspaper reports of daily high and low temperatures in various locations in the SoCalGas service territory. These were used to construct heating and cooling degree-day variables that were matched to the period of each gas bill. The last was a survey of a sample of participants and nonparticipants that gathered information on the respondents' demographic and dwelling characteristics.

We merged the information from these sources to construct a time series/cross sectional dataset of consumption for a sample of participants and nonparticipants spanning the period from twelve months prior to the 1997 program period through eleven months afterward (i.e., from September 1996 through November 1998). This dataset permitted the comparisons of changes in gas consumption for participants and nonparticipants from before to after the program treatment, controlling for changes in weather, and other non-program effects.

In the course of the analysis, several different equation specifications were estimated. All of the models specified average daily gas use as the dependent variable and a core set of explanatory variables. These core explanatory variables included a binary variable for participation, heating degree-days, several variables indicating equipment holdings, and demographic variables. The various specifications, which entailed adding and redefining variables to the core set of independent variables, are presented and discussed in Chapter 3.

In the final model specification selected to estimate the net HEF Program impacts, the program effects were represented by three distinct variables that capture different aspects of program participation. The first was a binary variable equal to one for participants in all billing periods, and zero for the nonparticipants. The second was a binary variable equal to one for all customers for all billing periods during 1998 and zero prior to 1998. The third was a binary variable equal to one for program participants after they received the audit report, and zero for them before. The variable was zero for nonparticipants in all periods. The first variable captured any underlying differences in average consumption between participants and nonparticipants, after controlling for demographics and dwelling characteristics. The second variable captured any underlying change in gas consumption in the last year. This change is interpreted as the result of any "naturally occurring" conservation (i.e., free ridership effects), as well as the effects of background economic variables on consumption. The last variable captured the average net change in gas consumption for participants from before to after receipt of the audit report, after netting out the naturally occurring conservation, and controlling for weather changes and other effects.

ORGANIZATION OF REPORT

The remainder of this report in organized as follows. Chapter 2 describes, in detail, the data used in the analysis. Issues regarding the sampling procedure and connected data preparation are outlined. The associated Appendices contain a detailed accounting of specific information when appropriate.

The methodology underlying the statistical analysis used to estimate the energy savings impact of SoCalGas' 1997 HEF Program is presented in Chapter 3. Additional potential and actual statistical problems are addressed and their solutions outlined when applicable.

The empirical results of the conditional demand analysis and the estimates of net and gross impacts on gas energy savings are reported in Chapter 3.

Chapter 2
DATA EXTRACTION, SAMPLING, AND SURVEY ADMNISTRATION

DATA EXTRACTION AND PREPARATION

The data used to estimate the HEF Program impacts were drawn from the following four sources:

- The HEF Program Tracking Database;
- SoCalGas' billing system;
- Local newspaper temperature readings at various locations in the SoCalGas service territory; and
- A telephone survey of a sample 1997 HEF participants and nonparticipants.

Initially SoCalGas provided a file of all 1997 HEF participants with continuous billing records from January 1996 through November 1997. This file contained information for 23,650 customers categorized as 1997 participants. In addition, SoCalGas provided the bills for a large random sample of nonparticipant residential customers, with continuous service since January 1992, who live in single-family residences. The length of service and dwelling type criteria matched those upon which SoCalGas had targeted its HEF marketing efforts. The size of this initial nonparticipant random sample was 90,182 customers.

The original datasets provided by SoCalGas were screened according to several criteria. The screening eliminated accounts that failed minimum data requirements for even a simple before-and-after comparison of the participant and control populations. The screening criteria consist of the following:

1. The sample was restricted to single-family residential accounts only. This means that the building code is restricted to 'A' (single-family detached dwelling) and that the rate code is either GR (residential) or GR-L (low-income residential). All nonparticipant accounts received from SoCalGas were already screened for this.

2. Participants with audit dates in January 1998 or missing audit dates were dropped (meaning the audit report had not been mailed back as of mid-January 1998), since we would not have a full year of post-audit billing data for them. One participant with an audit date of January 1997 was also eliminated at this point. We assume this was a typographical error. All remaining participants have audit dates in October, November, or December 1997.

3. Several participants and a few nonparticipants with service addresses outside of the SoCalGas service territory were eliminated (by zip code).

4. A few accounts with missing weather zones were screened out because weather zone is needed to stratify the sample.

5. Accounts that were inactive in 1996 were eliminated. These are accounts where the number of days billed in 1996 is zero, according to the data received from SoCalGas.

There are quite a few participants screened out according to this criterion. We presume these accounts were reactivated before the HEF surveys were sent out.

6. We required that the number of days billed in 1996 be more than 340. If the threshold number were changed here from 340 to 250, we would gain only 3 participants and no nonparticipants.

7. In the nonparticipant file received from SoCalGas, there are no accounts with a service connection date later than 1991. Only 1.5% of participants (unfiltered) has a service connection date later than 1991. This percentage drops to 0.7% if the threshold year is set at 1992, so we have eliminated those participants with a turn-on date greater than 1992 to make the two populations comparable.

8. We screened out one nonparticipant with a service connection date earlier than 1961. The 1997 HEF Program targeted high-usage customers in certain climate zones (3, 4 & 5). By way of restricting the range of usage of the nonparticipant sample to be comparable to that of the participant sample, customers with 1996 usage of less than 200 therms were eliminated, as well as those with usage greater than 6500 therms. This filter affects only nonparticipants, except for three participants whose use is significantly lower than the rest of the participant sample.

9. Several of the 1997 participants and nonparticipants have participated in a HEF Program in previous years. We eliminated only those whose previous audit year was 1996, since that effectively eliminates our "before" picture. After this final screening, there are still 2704 participants who participated in earlier years, and 3390 nonparticipants.[1]

Initially there are 23,650 participant observations and 90,182 nonparticipants. After the screening, 21,133 participant accounts and 86,438 nonparticipant accounts remained. Table 2-1 (page 2-3) documents the number of accounts that were eliminated at each point in the screening process.

Once we screened the original datasets and eliminated a small number of cases according to the criteria documented above, we stratified the remaining cases by weather zone, service connection date (turn-on year), and consumption level. The distribution of the participant and nonparticipant billing samples by these strata are presented in Table 2-2 (page 2-3). As the percentages indicate, the distributions of the participants and nonparticipants by turn-on year are fairly comparable. The distributions by weather zone and average gas consumption differ significantly, reflecting the targeting of the offer at households with above-average gas consumption, who tend to live in colder parts of the service territory.

[1] In the dataset that was sent to the survey subcontractor, these cases were also eliminated. As a result, none of the households who were surveyed had received an audit from SoCalGas since 1992.

Table 2-1
Customer Account Attrition Due to Screening

	Participant Account	*Nonparticipant Account*
Original Sample	**23,650**	**90,182**
Reason for Deletion		
Non-Single Family Dwelling	96	0
Nonresidential Account	14	0
Audit Date After December 1997	1,000	0
Address outside Service Territory	24	3
Missing Weather Zone Information	3	0
Inactive Account in 1996	859	4
Less than 340 Billing Days in 1996	10	0
Service Connection Date after 1992	18	0
Service Connection Date before 1961	0	1
Less than 200 Therms in 1996	3	2,978
More than 6500 Therms in 1996	0	20
Audited in 1996	490	738
Total Number of Deleted Accounts	2,517	3,744
# of Observations Remaining	21,133	86,438

Table 2-2
Billing Sample Distributions by Stratification Variables

		Participant Billing Sample	*Nonparticipant Billing Sample*
Turn-On Year			
1	Prior to 1980	46.0%	46.9%
2	1980 or Later	54.0%	53.1%
Weather Zone			
1	Mountain	0.7%	0.7%
2	Lower Desert	2.4%	2.2%
3	Coastal Strip	41.1%	19.8%
4	Upper Desert	16.5%	6.3%
5	Inland Valley	37.2%	41.2%
6	LA Basin	2.1%	29.7%
1996 Use-Per-Day			
1	Less Than 1.6 therms/day	34.4%	60.5%
2	1.6-2.0 therms/day	35.5%	19.1%
3	Greater than 2.0 therms/day	30.2%	20.4%

The combinations of turn-on date, weather zone, and average daily gas consumption yield a total of 36 strata. Some of these were merged for purposes of drawing the telephone survey sample because of the small number of customers in the certain strata. All of the observations in Weather Zone 1 were combined into a single stratum, and customers in Zones 2 and 6 were split into two strata only – those with less than 2 therms/day and with more than 2 therms/day. After these combinations, there were a total of 23 strata from which the telephone survey samples were drawn. The stratum definitions are presented in Table 2-3.

Table 2-3
Telephone Sample Distributions

Stratum	Turn on Year	Weather Zone	Therms per Year	*Participant Survey Sample*	*Nonparticipant Survey Sample*	*Target Percent*	*Target No.*	*Completed Nonparticipant Survey*	*Completed Participant Survey*
A	**All**	**1**	**All**	26	25	0.7	4	3	4
B	**All**	**2**	**< 2.0**	64	57	1.7	8	5	11
C	**All**	**2**	**> 2.0**	27	23	0.7	4	3	4
D	**Prior to '80**	**3**	**< 1.6**	306	289	8.3	41	39	50
E	**80 or Later**	**3**	**< 1.6**	215	202	5.8	29	27	35
F	**Prior to '80**	**3**	**1.6 to 2.0**	298	281	8.1	40	41	49
G	**80 or Later**	**3**	**1.6 to 2.0**	242	225	6.6	33	30	39
H	**Prior to '80**	**3**	**> 2.0**	294	260	8	40	36	42
I	**80 or Later**	**3**	**> 2.0**	247	212	6.7	34	34	36
J	**Prior to '80**	**4**	**< 1.6**	47	44	1.3	6	7	5
K	**80 or Later**	**4**	**< 1.6**	135	127	3.7	19	20	23
L	**Prior to '80**	**4**	**1.6 to 2.0**	61	55	1.7	8	9	14
M	**80 or Later**	**4**	**1.6 to 2.0**	153	142	4.1	21	27	22
N	**Prior to '80**	**4**	**> 2.0**	52	49	1.4	7	5	8
O	**80 or Later**	**4**	**> 2.0**	123	105	3.3	17	20	15
P	**Prior to '80**	**5**	**< 1.6**	138	131	3.7	18	18	24
Q	**80 or Later**	**5**	**< 1.6**	345	328	9.3	46	51	48
R	**Prior to '80**	**5**	**1.6 to 2.0**	132	124	3.6	18	19	28
S	**80 or Later**	**5**	**1.6 to 2.0**	350	314	9.5	48	44	51
T	**Prior to '80**	**5**	**> 2.0**	112	106	3	15	17	15
U	**80 or Later**	**5**	**> 2.0**	242	229	6.6	33	36	32
V	**All**	**6**	**< 2.0**	57	56	1.5	7	8	10
W	**All**	**6**	**> 2.0**	25	24	0.7	4	4	4
	Total			3691	3408	100	500	503	569

The telephone survey samples were drawn from the stratified participant and nonparticipant billing datasets. The participant survey sample was drawn proportionately from the billing dataset. The nonparticipant survey sample was drawn so that its distribution by stratum closely matched those of the participant sample.

Samples of 3,691 participants and 3,408 nonparticipants were sent to the telephone survey subcontractor. The contractor was instructed to conduct at least 500 complete surveys of the participants and nonparticipants. Completion quotas were set by stratum so that the distributions of completed surveys would track those of the samples sent to the subcontractor.

SURVEY ADMINISTRATION

The survey implementation subcontractor randomized the combined participant and nonparticipant gross samples and broke them up into subsets of approximately 1000 cases. The first subset loaded the file into its Computer Aided Telephone Interview (CATI) system. The second subset was only used after an attempt had been made to contact all of the customers in the first subset, as a means to minimize non-response bias. The CATI system ensured a high quality survey administration, with randomized calling patterns, quotas in each sample stratum, tracking and documentation of the sample disposition, automatic skip patterns execution, and verification of responses according to various criteria.

The CATI coded survey instrument was pre-tested on a small sample of households and modified based on the experience of respondents with the clarity of the questions and the flow of the interview. There were very few changes to the survey script because a similar one had been used successfully in the evaluation of the market effects of same program (Home Energy Fitness Program Market Effects Evaluation, May 1998) for SoCalGas.

The telephone surveys were administered on weekday evenings and all day Saturdays by trained telephone interviewers. The interviewers made up to six attempts to contact households in the sample.

The disposition of the combined sample is summarized in Table 2-4.

Table 2-4
Disposition of Telephone Survey Sample

	Count	Percent
Sample	7,099	
Total Attempted Contacts	2,610	36.8%
Total Phone Calls Placed	5,738	
Calls Per Attempt	2.20	
Non-contact		
Not contacted	867	
Not In Service	120	4.6%
Busy	25	1.0%
No Answer	193	7.4%
Answering Machine / Voice Mail	274	10.5%
Wrong Number	51	2.0%
Other phone problems	9	0.3%
Contact		
Requested callback	454	17.4%
Refused	370	14.2%
Unable to provide information	28	1.1%
Language Barrier (other than Spanish)	11	0.4%
Complete Interviews	1,075	41.2%

We calculated the reliability of the sample based on the criteria set forth in the M&E Protocols. These criteria state that, "a sample must be randomly drawn and be sufficiently large to achieve a minimum precision of plus/minus 10% at the 90% confidence level, based on total annual energy use." (M&E Protocols, Table 5, p.14). We computed the sample reliability for estimating the 1996 annual gas consumption of the participants and nonparticipants, respectively. In order to accomplish this, we calculated the average 1996 gas consumption per account and its standard deviation for each stratum in the screened participant and nonparticipant billing datasets. Based on these population values, we computed the variances of the mean sample estimates of gas consumption in each stratum, using the number of completed surveys in each stratum. These were used, in turn, to calculate the overall standard error of the mean estimate of average daily gas consumption from the completed survey sample. The population average daily use and standard errors of the sample means are presented in Table 2-5. As they show, the standard error for the participants group is .022, which is approximately one-ninth of the 10% "margin of error" for the sample.

Table 2-5
Reliability of Survey Sample

	Nonparticipants					*Participants*				
Stratum	***Number in Population***	***Mean Daily Use***	***Standard Deviation***	***Number in Sample***	***Standard Error***	***Number in Population***	***Mean Daily Use***	***Standard Deviation***	***Number in Sample***	***Standard Error***
A	145	1.99	0.73	3	0.417	139	2.05	0.63	4	0.310
B	361	1.34	0.46	5	0.204	364	1.6	0.19	11	0.056
C	189	3.39	2.15	3	1.231	152	3.02	1.88	4	0.928
D	1863	1.16	0.27	39	0.043	1716	1.47	0.08	50	0.011
E	1143	1.13	0.28	27	0.053	1130	1.48	0.07	35	0.012
F	1682	1.78	0.11	41	0.017	1682	1.77	0.11	49	0.015
G	1522	1.78	0.11	30	0.020	1258	1.78	0.11	39	0.017
H	1796	2.88	1.46	36	0.241	1608	2.77	1.17	42	0.178
I	927	2.97	1.5	34	0.252	1289	2.78	1.03	36	0.169
J	383	1.21	0.25	7	0.094	411	1.47	0.08	5	0.036
K	770	1.17	0.26	20	0.057	715	1.47	0.08	23	0.016
L	425	1.79	0.12	9	0.040	499	1.77	0.11	14	0.029
M	679	1.78	0.11	27	0.021	815	1.78	0.11	22	0.023
N	475	2.55	0.56	5	0.249	394	2.52	0.63	8	0.220
O	599	2.51	0.73	20	0.160	657	2.48	0.56	15	0.143
P	1181	1.18	0.26	18	0.061	1087	1.46	0.09	24	0.018
Q	2033	1.15	0.27	51	0.037	1813	1.46	0.09	48	0.013
R	1135	1.78	0.11	19	0.025	1058	1.77	0.11	28	0.021
S	2027	1.78	0.11	44	0.016	1825	1.77	0.11	51	0.015
T	803	2.8	1.03	17	0.247	828	2.54	0.69	15	0.177
U	1436	2.73	1.01	36	0.166	1256	2.51	0.62	32	0.108
V	361	1.45	0.38	8	0.133	306	1.61	0.2	10	0.062
W	153	2.88	1.48	4	0.730	131	2.83	1.09	4	0.537
e.	22088	1.86	0.96	503	0.031	21133	1.93	0.72	569	0.022

For the nonparticipant sample, the standard error is one-sixth of the 10% tolerance level. These relative magnitudes mean that the reliability of the sample greatly exceeds the minimum levels required by the M&E Protocols. In fact, they exceed the 99% confidence level for a 5% tolerance level.

SURVEY QUESTIONS AND SUMMARY RESULTS

The telephone survey asked a series of questions about the respondents' gas equipment holdings and operation, recall of the HEF audit (participants only), installation of measures and adoption of conservation practices in the past two years, and socioeconomic and dwelling characteristics. A copy of the telephone survey questionnaire is presented in Appendix B.

A summary of key information from the survey is presented in Table 2-6. A complete listing of the frequencies of the telephone survey responses appears in Appendix C.

Table 2-6
Summary of Selected Survey Frequencies

	Participants	***Nonparticipants***
Member of Household 65+ Years Old?	0.523	0.459
Household Member a College Graduate?	0.504	0.497
Annual Income Greater than $50K?	0.263	0.276
Home Built Since 1960?	0.595	0.531
Four or More Bedrooms?	0.307	0.316
Gas Space Heating?	0.988	0.974
Good Attic Insulation?	0.529	0.506
Setback Thermostat?	0.400	0.339
Change Furnace Filter 3+ times per Year?	0.361	0.365
Turn Off Pilot during Summer?	0.291	0.298
Replaced Furnace in Past 2 Years?	0.117	0.098
Gas Water Heater?	0.972	0.976
Replaced Water Heater in Past 2 Years?	0.226	0.173
Own Gas Clothes Dryer?	0.786	0.758
Gas Heated Pool or Spa?	0.206	0.178
Recall Receiving HEF Audit?	0.161	NA

The responses to the demographic questions in the survey indicate that the participants and nonparticipants are fairly similar with respect to education and income. The differences between the two groups are not statistically significant.

The age distributions of the two groups appear to be significantly different. 52% of participant households contain one or more member over 65 years of age, versus 46% for the nonparticipant group. The six-percent difference is statistically significant at the 90% confidence level.

With respect to dwelling size as measured by number of bedrooms, the two groups are relatively comparable. However, the participant's homes tend to be somewhat newer than those of the nonparticipants. Recall that the sample was stratified by service connection date (before 1980 and 1980 or later), which should control for differences in dwelling age, to some degree. Nonetheless, the participants have a significantly larger percent of homes built after 1960.

The gas equipment ownership and recent replacement patterns of participants and nonparticipants are generally comparable. The saturations of gas space and water heating for the two groups are nearly 100%, as would be expected for households in Southern California with gas service. The saturation of gas dryers is also relatively high (79% for participants and 76% for nonparticipants), and the saturations of "good" insulation are moderate (around 50%). Participants have a moderately higher saturation of setback thermostats than nonparticipants (40% versus 34%). They also have slightly higher ownership rates of pools and spas (21% versus 18%), although the difference is not statistically significant.

Recent equipment replacement patterns differ moderately between the two groups – possibly due to the effects of the HEF Program. Approximately 2% more participants have replaced their gas furnaces in the past two years (12% versus 10%), while 5% more have replaced their water heaters (23% versus 17%).

Other key conservation practices appear comparable for the two groups. The rates at which participants and nonparticipants replace their furnace filters are similar, as are the percentages who turn off their pilot lights during the summer.

One of the most noticeable results of the survey is the rate at which participants recall the receipt of the HEF audit. Only 16% of the participants said that they recalled receiving the HEF report in the past two years. While this low recall rate is striking, it is consistent with the results of previous evaluations of the HEF Program (1994 impact evaluation and 1998 market effects evaluation). As a practical matter, the low recall rate makes it impossible to directly determine through survey questions whether the audit recommendations caused participants to adopt any measures.

PREPARATION OF THE ANALYSIS DATABASE

As the last step in preparing the data for the regression analysis whose results are reported in the following chapter, the survey responses were merged with the monthly billing data and weather variables corresponding to the days covered by the gas consumption data. The billing, survey, and weather data were "stacked" so that each month of consumption represents a single record. The stacked dataset was "squared" so that there were exactly 27 records for each household in the sample.[2] These 27 records correspond to the 27 months from September 1996 through November 1998 – the period spanning one year before the receipt of audits through eleven months after the deliver of the last audit report (the latest available bills).

[2] This last step of "squaring" the dataset was required by the statistical procedure used in the analysis.

In the course of preparation of the analysis dataset, a few records were eliminated due to problems with the billing data. The first problem consisted of the presence of gaps between the bills. For some customers, the end date of one billing cycle fell several days before the start date of the following bill. In these cases (a total of 12), all records for the customer were deleted.

The second problem involved cases where the customer had fewer than 27 billing months of consumption. The statistical procedure used to estimate the parameters of the model requires that the database contain the same number of observations per household. In those instances where there were less than 27 observations, we deleted all records for the household. This only affected 5 households. A summary of the data attrition for participants and nonparticipants is presented in Table 2-7.

Table 2-7
Data Attrition Due to Billing Data Problems

	Participants	***Nonparticipants***
Original Number of Households	569	503
Households with Data Gaps	5	7
Households with < 27 Months	1	4
Households in Analysis Dataset	563	492

Chapter 3
METHODOLOGY AND RESULTS

METHODOLOGY

Regression Based Demand Analysis

We performed a regression analysis on the merged billing, weather, and survey data to estimate the natural gas savings attributable to the 1997 HEF Program. The objective of the regression analysis was to estimate the net impact of the HEF Program, after controlling for other factors that influence natural gas consumption. In addition, the analysis quantified differences in participant and nonparticipant natural gas usage and differences in usage due to underlying economic and environmental trends.

The general form of the specification of the conditional demand relationship is:

$$Y_{it}=\Sigma\beta_k X_{kt} + \varepsilon_{it}$$

where

i = 1,2,3,...,I (I=total number of accounts);
t = 1,2,3,…,T (number of billing cycles);
Y_{it} = average natural gas use per day in billing cycle t by customer i;
X_{kt} = the k explanatory variable defined in Table 3-1;
β_k = the k-th coefficient;
ε_{it} = random component representing unobserved factors affecting consumption;

A linear specification was chosen because the relationship between total gas consumption and the ownership of different end uses is arithmetically linear.

The parameters of a basic model specification were estimated using statistical regression procedures that test and correct for the presence of serially correlated errors and heteroskedasticity, when either is found. The tests indicated that the errors in the models were both serially correlated and heteroskedastic. A detailed discussion of the techniques used to test and correct for these problems is presented later in this chapter.

Table 3-1
Variable Definitions and Mean Values
(Page 1 of 2)

Variable	Definition	Mean	
		Partic	Nonpart
A65P	1 if anyone in the household is 65 years or older, else 0	0.7708	0.6261
A65CDSF	A65P*CDSF	1.4825	1.1227
A65HDSF	A65P*HDSF	3.2746	2.7421
CDD70	number of base 70 cooling degree-days during billing period	51.071	51.046
CDSF	(CDD70/NDAYS)*SQFTI	3.1476	3.0417
CDSFQ4	CDSF*Y98Q4	0.0510	0.0519
CKPILOT	1 if the gas stove has a pilot light, else 0	0.2149	0.2602
COVER	POOLCVR or SPACVR	0.1332	0.1199
GASCOOK	1 if any of the cooking equipment uses gas, else 0	0.8135	0.8455
GASDRYER	1 if the home includes a gas clothes dryer, else 0	0.7893	0.7664
GASHEAT	1 if the primary heating fuel is gas, else 0	0.9876	0.9736
GASPOOL	1 if the home has a gas-heated swimming pool, else 0	0.1191	0.0854
GASSPA	1 if the customer owns a gas-heated spa or jacuzzi, else 0	0.1776	0.1646
GASWTRHT	1 if the water-heating fuel is gas, else 0	0.9838	0.9796
HDD65	number of base 65 heating degree-days during billing period	120.44	123.16
HDSF	HDD65*SQFTI	7.0144	7.1000
HDSFQ4	HDSF*Y98Q4	0.6190	0.6510
KIDS	1 if anyone in the household is under 6 years old, else 0	0.0777	0.0962
NDAYS	number of days in the billing period	30.460	30.468
PARTIC	1 if the customer was a 1997 participant, else 0	1.0000	0.0000
PCDSF	PARTIC*CDSF	3.1476	0.0000
PHCDSF	POSTHEF*CDSF	1.2962	0.0000
PHDSF	PARTIC*HDSF	7.0144	0.0000
PHHDSF	POSTHEF*HDSF	3.6584	0.0000
POOLCVR	1 if the gas-heated swimming pool has a cover, else 0	0.0569	0.0407
POOLSPA	GASPOOL or GASSPA	0.2043	0.1809
POSTHEF	1 if the entire billing period post-dates the audit, else 0	0.4464	0.0000
POSTHF98	POSTHEF*Y1998	0.4255	0.0000
POSTQ1	POSTHEF*Y98Q1	0.1035	0.0000
POSTQ2	POSTHEF*Y98Q2	0.1136	0.0000
POSTQ3	POSTHEF*Y98Q3	0.1141	0.0000
POSTQ4	POSTHEF*Y98Q4	0.0943	0.0000
POSTQ3B	POSTHEF*Y98Q3B	0.1452	0.0000
SPACVR	1 if the gas-heated spa/jacuzzi has a cover, else 0	0.1172	0.1098
SQFT	gas-heated square feet in home	1875.5	1816.3
SQFTI	SQFT, with imputed values where responses were missing	1856.4	1814.1

Table 3-1
Variable Definitions and Mean Values
(Page 2 of 2)

Variable	Definition	Mean	
		Partic	**NonPart**
TCONST	intercept term	1.0000	1.0000
W14CDSF	WZ14*CDSF	0.6474	0.7470
W14HDSF	WZ14*HDSF	2.1544	2.4358
W26CDSF	WZ26*CDSF	0.5531	0.2941
W26HDSF	WZ26*HDSF	0.2452	0.1454
W3CDSF	WZ3*CDSF	0.4609	0.4195
W3HDSF	WZ3*HDSF	2.6304	2.3958
W5CDSF	WZ4*CDSF	1.4862	1.5812
W5HDSF	WZ4*HDSF	1.9844	2.1229
WZ14	1 if the weather zone is Mountain or Upper Desert, else 0	0.1599	0.1768
WZ26	1 if the weather zone is Lower Desert or LA Basin, else 0	0.0497	0.0407
WZ3	1 if the weather zone is Coastal Strip, else 0	0.4441	0.4146
WZ5	1 if the weather zone is Inland Valley, else 0	0.3464	0.3679
Y1998	1 if the read date occurs in 1998, else 0	0.4294	0.4320
Y98Q1	read date is in period Jan-Mar 1998	0.1074	0.1077
Y98Q2	read date is in period Apr-Jun 1998	0.1136	0.1132
Y98Q3	read date is in period Jul-Sep 1998	0.1141	0.1146
Y98Q4	read date is in period Oct-Dec 1998	0.0943	0.0964
Y98Q3B	Y98Q3 or Y98Q4, 25-period runs only	0.1378	0.1405

REGRESSION MODEL SPECIFICATION AND VARIABLE DEFINITIONS

The definitions for the variables used in the different model specifications are presented in Table 3-1. The basic variables generally fall into five categories. The first are binary variables representing the ownership of a given type of gas equipment or a measure affecting gas consumption. For example, the variable GASCOOK is defined as 1 if the household has a gas stove or oven, zero otherwise. Other equipment variables include CKPILOT (gas cooking pilot), GASDRYER, GASHEAT, GASPOOL, GASSPA, GASWTRHT, and POOLCVR.

The second category of variables represents the household and dwelling characteristics that may affect gas consumption. For example, we expect the size of the dwelling, represented by the square feet of the dwelling, to affect gas consumption. Other household/dwelling variables include binary variables representing young children and persons over 64 in the household (KIDS and A65P, respectively).

The third category is weather variables. These are average daily heating and cooling degree-days for the billing period (HDD65 and CDD70, respectively).

The fourth category is geographical variables. These are all binary variables representing the different weather zones in the SoCalGas service territory.

The last category contains the key variables that capture the effects of the HEF Program, as well as any "background" factors that are contemporaneous with the program impacts. These are PARTIC, Y1998, and POSTHF98. PARTIC is defined as one for participants and zero for nonparticipants. It captures any underlying difference in gas consumption between the two groups after controlling for other observable variables. Y1998 captures any underlying change in gas usage in the population for the calendar year of 1998. It is defined as 1 for both participants and nonparticipants for all billing periods from January 1998 and afterward, and zero prior to January 1998. The coefficient of Y1998 represents the combined effects of economic and environmental trends, as well as the impacts of any "naturally" occurring conservation actions, on the average daily gas usage that occurred during the last year.

POSTHF98 is defined as one for participants in 1998 after they received the HEF audit report, and zero in preceding months. When POSTHF98 enters the specification along with Y1998, it is interpreted as net impact of the HEF Program, after controlling for any naturally occurring conservation and trends in consumption (picked up by Y1998).

Several compound variables are specified in different regression models using the basic variables. These are defined as the algebraic products of the basic variables, and they are intended to capture certain interaction effects in gas consumption patterns. For example, the effect of weather on gas consumption is expected to depend itself on the size of the dwelling. This is represented by the variable HDSF, a term that is the product of heating degree-days and dwelling square feet. Several such compound variables are used in different model specifications to explore these types of interactive effects.

No attempt was made in any of the models to estimate the program impacts by end use

category. There were two primary reasons for this. First, the program tracking did not record what measures and actions were recommended to each recipient of the audit report. While the survey asked respondents about specific actions they took, the response rates to these questions were very low. As a result, there was no way to identify what end uses could have been affected by the recommendations with any degree of reliability.

Second, almost all of the customers in the sample had both space and water heating. If binary variables representing these end uses interacting with program participation variables were included in the specification, we expected that the program effects would be confounded by the high level of collinearity among the end use variables and the constant term in the model.

In the model specifications where the key program participation variable, POSTHF98, interacts with heating degree-days (discussed below), one might argue that the effect of the program on space heating consumption is captured by this weather sensitive variables. However, water heating has a seasonal use pattern whose effect would be partially picked up by this variable. We reject the argument that the weather sensitive portion of any program impacts can be attributed exclusively to space heating.

PRESENTATION AND DISCUSSION OF PARAMETER ESTIMATES

In the initial model specification, daily gas usage was regressed on the following set of variables:

- Geographical indicators (WZ14, WZ26, WZ5);
- Gas equipment holdings (GASHEAT, GASWTRHT, GASDRYER, GASCOOK, CKPILOT, GASPOOL, POOLCVR, GASSPA, SPACVR)
- Heating and cooling degree days interacting with square feet (HDSF, CDSF);
- The variables capturing the program effects and the post audit trends, entering alone and interactively with square feet and degree-days (PARTIC, HDSF, PHDSF, PHHDSF, CDSF, PCDSF, PHCDSF, Y1998, and POSTHF98).

The parameters of the initial model specification were estimated using ordinary least squares regression techniques, and standard tests were applied to determine whether the error structure is autocorrelated and/or heteroskedastic. The results of these tests indicated the presence of both. As a result, the model parameters and standard errors were re-estimated using procedures, described below, that correct for these statistical problems. All of the results reported here were obtained using these techniques.

The results of Model 1A are reported in Table 3-2. While the overall explanatory power of the model is good (R^2 = .57), there are some problems with the values of some of the parameter estimates and their statistical performance. The values of most of the parameters for the equipment holdings are either the wrong sign (negative), or statistically insignificant. The coefficients for GASHEAT and GASWTRHT are negative. The coefficients for GASDRYER, GASCOOK, GASPOOL, and GASSPA are all statistically insignificant.

The coefficients of the weather variables interacting with square feet (HDSF, CDSF) are the correct sign and statistically significant. The coefficients of the heating degree variable for participants is positive and statistically significant, indicating that this group's gas use is more sensitive to temperature than nonparticipants.

Table 3-2
Model 1A: All Appliances

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.4472	7.7614
WZ14	-0.7609	-16.5211
WZ26	0.2844	3.2766
WZ5	-0.0772	-2.2806
PARTIC	0.0372	0.8420
GASHEAT	-0.2725	-1.4967
GASWTRHT	-0.0632	-0.5587
GASDRYER	0.0198	0.4526
GASCOOK	0.0130	0.3081
CKPILOT	0.1319	3.6236
GASPOOL	0.0845	0.7825
POOLCVR	0.0004	0.0026
GASSPA	0.0401	0.4151
SPACVR	-0.2202	-2.1534
HDSF	0.1029	52.6844
PHDSF	0.0149	5.3592
PHHDSF	-0.0027	-1.1445
CDSF	-0.0077	-5.1911
PCDSF	0.0003	0.1391
PHCDSF	-0.0022	-0.9327
Y1998	0.2669	7.8966
POSTHF98	-0.0419	-0.8805
Adjusted R-Squared:		0.5711
Number of Observations:		28080
Number of Accounts:		1040
Billing Records Per Account:		27
Autoregressive Parameter		-0.71752
t-statistic		-172.55

The binary indicator for participants (PARTIC) shows that they have slightly greater overall gas use than nonparticipants do. The magnitude of Y1998 indicates that there has been an overall trend of increasing gas consumption in 1998 relative to the pre-audit period.

The variables that capture the net program effects are POSTHF98, PHHDSF, and PHCDSF. The values of the parameters for these variables are negative, indicating that the

program participation has reduced gas consumption after controlling for other factors. However, the statistical significance of these estimates is low. If the point estimates of the parameters are used to estimate net program savings, the value of the savings estimate is 25.8 therms per participant per year.

Model 1B (Table 3-3) drops GASHEAT and GASWTRHT from the specification. The reasoning for this is that almost all of the respondents have gas space and water heating. The binary indicators are highly collinear with the constant term, making it impossible to identify the effects of these end uses separately. The omission of these variables has little effect on the estimates for the remaining model parameters, however.

Table 3-3
Model 1B: No GASHEAT or GASWTRHT

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.1314	21.0465
WZ14	-0.7608	-16.6410
WZ26	0.2894	3.3524
WZ5	-0.0716	-2.1053
PARTIC	0.0364	0.8410
GASDRYER	0.0114	0.2650
GASCOOK	0.0022	0.0557
CKPILOT	0.1313	3.6260
GASPOOL	0.0841	0.7859
POOLCVR	0.0088	0.0574
GASSPA	0.0277	0.2867
SPACVR	-0.2173	-2.1385
HDSF	0.1031	52.8700
PHDSF	0.0148	5.3547
PHHDSF	-0.0027	-1.1185
CDSF	-0.0077	-5.1638
PCDSF	0.0002	0.1017
PHCDSF	-0.0022	-0.9282
Y1998	0.2661	7.8987
POSTHF98	-0.0410	-0.8648
Adjusted R-Squared:		0.5711
Number of Observations:		28080
Number of Accounts:		1040
Billing Records Per Account:		27
Autoregressive Parameter		-0.71766
t-statistic		-173.29

Model 1C (Table 3-4) combines the variables for gas-heated pools and spas into a single variable. The rationale for this is that many people have pool-spa combinations, making identification of the separate effects of these end uses difficult. This produces a reasonable and statistically significant estimate of the value for this parameter. At the same time, the parameter for COVER is negative and greater in absolute magnitude than POOLSPA. This means that use of a pool or spa cover is a proxy for various gas saving measures and behavior that more than offset the average net contribution of gas pool/spa heating on overall consumption.

Table 3-4
Model 1C: POOL and SPA Combined

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.1199	20.6117
WZ14	-0.7604	-16.6090
WZ26	0.2923	3.4151
WZ5	-0.0682	-1.9733
PARTIC	0.0388	0.8887
GASDRYER	0.0027	0.0618
GASCOOK	0.0125	0.3096
CKPILOT	0.1359	3.7784
POOLSPA	0.1562	2.1257
COVER	-0.2739	-3.1358
HDSF	0.1030	52.8026
PHDSF	0.0149	5.3555
PHHDSF	-0.0026	-1.1130
CDSF	-0.0077	-5.2072
PCDSF	0.0003	0.1199
PHCDSF	-0.0022	-0.9266
Y1998	0.2667	7.9181
POSTHF98	-0.0416	-0.8783
Adjusted R-Squared:		0.5710
Number of Observations:		28080
Number of Accounts:		1040
Billing Records Per Account:		27
Autoregressive Parameter		-0.71784
t-statistic		-173.38

A second set of model specifications examines the effects of demographic variables on gas consumption. Model 2A (Table 3-5) adds a binary variable, KIDS, equal to one if there are children under six in the household. In addition, the variables for household members over 65 are specified interactively with heating and cooling degree-days. This captures the hypothesis that weather sensitive requirements for the elderly are greater than the requirements for younger households. The parameter estimates for all three variables are reasonable and statistically significant.

Table 3-5
Model 2A: Age Variables Added

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.1129	20.7640
WZ14	-0.8127	-18.3429
WZ26	0.1552	2.2096
WZ5	-0.0965	-2.9166
PARTIC	0.0293	0.6968
GASDRYER	0.0471	1.2199
GASCOOK	-0.0061	-0.1500
CKPILOT	0.1373	3.7688
POOLSPA	0.1962	2.6161
COVER	-0.3397	-4.0283
KIDS	0.1487	2.5737
HDSF	0.0960	42.4401
A65HDSF	0.0182	7.7349
PHDSF	0.0139	4.9203
PHHDSF	-0.0011	-0.4758
CDSF	-0.0060	-3.5692
A65CDSF	-0.0044	-2.5165
PCDSF	0.0002	0.0969
PHCDSF	-0.0005	-0.2271
Y1998	0.2518	7.8112
POSTHF98	-0.0474	-1.0145
Adjusted R-Squared:		0.5903
Number of Observations:		26703
Number of Accounts:		989
Billing Records Per Account:		27
Autoregressive Parameter		-0.71320
t-statistic		-166.20

Model 2B (Table 3-6) adds the variable SQFTI. In earlier specifications, square feet entered interactively with heating and cooling degree-days. This additional variable reflects the hypothesis that non-weather related gas consumption may depend on the size of the dwelling. The parameters estimate is negative, which is counter to expectations. Given its significance, however, it is left in subsequent specifications.

Table 3-6
Model 2B: SQFTI Added

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.3934	17.9698
WZ14	-0.8381	-18.7132
WZ26	0.1721	2.4295
WZ5	-0.1079	-3.2651
PARTIC	0.0307	0.7313
GASDRYER	0.0604	1.5658
GASCOOK	-0.0374	-0.9339
CKPILOT	0.1169	3.1719
POOLSPA	0.2740	3.7762
COVER	-0.3550	-4.2569
SQFTI	-0.0002	-5.2071
KIDS	0.1541	2.6577
HDSF	0.0971	43.4832
A65HDSF	0.0181	7.6986
PHDSF	0.0137	4.8661
PHHDSF	-0.0009	-0.3950
CDSF	-0.0050	-2.9448
A65CDSF	-0.0048	-2.7248
PCDSF	0.0005	0.1982
PHCDSF	-0.0011	-0.4810
Y1998	0.2458	7.6185
POSTHF98	-0.0471	-1.0109
Adjusted R-Squared:		0.5915
Number of Observations:		26703
Number of Accounts:		989
Billing Records Per Account:		27
Autoregressive Parameter		-0.71255
t-statistic		-165.88

The next few models examine geographical variations in gas consumption in greater detail. Model 3A (Table 3-7) allows the weather sensitivity (both heating and cooling) to vary by zone. The coefficients for W14HDSF and W14CDSF are significantly different from those for the other weather ones. Based on this, it appears that houses in these areas are built significantly tighter than homes in the other parts of the SoCalGas service territory.

Table 3-7
Model 3A: DD Terms by Zone

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.3453	17.9022
WZ14	-0.4478	-9.1725
WZ26	0.3279	4.2948
WZ5	-0.0555	-1.5734
PARTIC	0.0387	0.9535
GASDRYER	0.0541	1.4881
GASCOOK	-0.0425	-1.1205
CKPILOT	0.1156	3.2880
POOLSPA	0.2605	3.6999
COVER	-0.3364	-4.1673
SQFTI	-0.0002	-6.7160
KIDS	0.1668	3.0198
W14HDSF	0.0752	27.9198
W26HDSF	0.1107	17.7733
W3HDSF	0.1281	42.9272
W5HDSF	0.1168	45.6888
A65HDSF	0.0184	8.2662
PHDSF	0.0112	4.2664
PHHDSF	-0.0029	-1.2296
W14CDSF	0.0111	3.6914
W26CDSF	-0.0111	-4.3262
W3CDSF	-0.0184	-5.1475
W5CDSF	-0.0093	-5.7800
A65CDSF	-0.0049	-2.8985
PCDSF	0.0025	1.1010
PHCDSF	-0.0020	-0.8808
Y1998	0.1762	5.6370
POSTHF98	-0.0292	-0.6445
Adjusted R-Squared:		0.6121
Number of Observations:		26703
Number of Accounts:		989
Billing Records Per Account:		27
Autoregressive Parameter		-0.70654
t-statistic		-163.06

Model 3B (Table 3-8) only allows the weather sensitivity to differ in Weather Zones 1 and 4, while the rest are constrained to be the same. These variables are retained in later specifications.

Table 3-8
Model 3B: DD Terms by WZ14 Only

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.3705	18.3776
WZ14	-0.4757	-9.9088
WZ26	0.2445	3.5417
WZ5	-0.1094	-3.4172
PARTIC	0.0361	0.8921
GASDRYER	0.0545	1.4981
GASCOOK	-0.0419	-1.1041
CKPILOT	0.1172	3.3213
POOLSPA	0.2595	3.6857
COVER	-0.3350	-4.1502
SQFTI	-0.0002	-6.6786
KIDS	0.1673	3.0186
HDSF	0.1215	54.3654
W14HDSF	-0.0467	-18.6450
A65HDSF	0.0190	8.5437
PHDSF	0.0113	4.3034
PHHDSF	-0.0028	-1.1869
CDSF	-0.0096	-6.1326
W14CDSF	0.0210	8.4624
A65CDSF	-0.0051	-2.9783
PCDSF	0.0030	1.2846
PHCDSF	-0.0032	-1.4136
Y1998	0.1781	5.6766
POSTHF98	-0.0242	-0.5349
Adjusted R-Squared:		0.6115
Number of Observations:		26703
Number of Accounts:		989
Billing Records Per Account:		27
Autoregressive Parameter		-0.70665
t-statistic		-163.13

The last set of models explores variations in savings over time. This is accomplished by allowing the trend in gas consumption and the program effects in 1998 to vary by calendar quarter. Model 4A (Table 3-9) includes separate binary variables for the four quarters of 1998 and for POSTHF98. The results show a striking variation in savings. The underlying trend in gas consumption shows a significant increase in the first quarter of 1998. This may be a result of the extreme “El Niño” impacts on gas consumption that are not fully reflected in the heating degree-day variable. At the same time, the non-weather sensitive program impacts increase substantially in the first and second quarter while they turn negative (i.e., the program increases

gas consumption) in the fourth quarter and cooling degree-day dependent effects become insignificant. If one includes the fourth quarter (negative) impacts, the net HEF Program impacts fall to approximately 17 Therms per year.

Table 3-9
Model 4A: Quartered POST

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.4044	18.8500
WZ14	-0.4705	-9.8723
WZ26	0.2378	3.4570
WZ5	-0.1103	-3.4803
PARTIC	0.0219	0.5273
GASDRYER	0.0550	1.5284
GASCOOK	-0.0413	-1.0973
CKPILOT	0.1168	3.3436
POOLSPA	0.2601	3.7375
COVER	-0.3342	-4.1915
SQFTI	-0.0002	-6.6835
KIDS	0.1694	3.0726
HDSF	0.1212	54.1833
W14HDSF	-0.0468	-18.7373
A65HDSF	0.0188	8.5312
PHDSF	0.0112	4.2900
PHHDSF	-0.0029	-1.2055
CDSF	-0.0091	-5.1456
W14CDSF	0.0202	8.0631
A65CDSF	-0.0052	-3.0223
PCDSF	0.0019	0.7728
PHCDSF	0.0009	0.3654
Y98Q1	0.2826	8.0080
Y98Q2	0.0164	0.4439
Y98Q3	0.0511	1.2640
Y98Q4	0.0968	2.0405
POSTQ1	-0.0639	-1.2223
POSTQ2	-0.0645	-1.1860
POSTQ3	-0.0233	-0.4155
POSTQ4	0.0676	1.0883
Adjusted R-Squared:		0.6158
Number of Observations:		26703
Number of Accounts:		989
Billing Records Per Account:		27
Autoregressive Parameter		-0.70354
t-statistic		-161.68

In an effort to clarify what is happening in the fourth quarter, Model 4B (Table 3-10) adds separate weather sensitive terms for the last two months. The results do not shed any light on why the effects are so different during this period. One possible explanation for the variations in results for the fourth quarter is due to the recent nature of the billing data. Often, billing data initially contains a moderate number of errors when it is first produced. These errors are corrected in due course through adjustments in the billing system. If the data covers a very recent period, then there is insufficient time for these corrections to be made.

Table 3-10
Model 4B: 4th Quarter DD Terms

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.4068	18.9543
WZ14	-0.4701	-9.8006
WZ26	0.2426	3.5235
WZ5	-0.1093	-3.4553
PARTIC	0.0210	0.5039
GASDRYER	0.0549	1.5250
GASCOOK	-0.0413	-1.0987
CKPILOT	0.1167	3.3431
POOLSPA	0.2601	3.7381
COVER	-0.3342	-4.1931
SQFTI	-0.0002	-6.7589
KIDS	0.1692	3.0725
HDSF	0.1209	52.3645
W14HDSF	-0.0467	-18.8357
HDSFQ4	0.0006	0.1489
A65HDSF	0.0188	8.5203
PHDSF	0.0114	4.3949
PHHDSF	-0.0031	-1.2933
CDSF	-0.0092	-5.1887
W14CDSF	0.0203	8.1159
CDSFQ4	-0.0147	-3.2051
A65CDSF	-0.0052	-3.0597
PCDSF	0.0020	0.8215
PHCDSF	0.0007	0.2576
Y98Q1	0.2820	7.9964
Y98Q2	0.0124	0.3322
Y98Q3	0.0427	1.0184
Y98Q4	0.0996	1.8592
POSTQ1	-0.0620	-1.1784
POSTQ2	-0.0622	-1.1400
POSTQ3	-0.0198	-0.3520
POSTQ4	0.0690	1.1106
Adjusted R-Squared:		0.6159
Number of Observations:		26703
Number of Accounts:		989
Billing Records Per Account:		27
Autoregressive Parameter		-0.70348
t-statistic		-161.65

Model 4C (Table 3-11) drops the last two months from the analysis dataset. The overall results improve dramatically, both in terms of plausibility and statistical performance. The parameter estimates for the equipment variables, the demographics, and the basic weather variables are generally reasonable in magnitude and statistically significant. The parameters of the variables capturing the program effects are also reasonable in magnitude, but they are only moderately significant from a statistical standpoint.

Table 3-11
Model 4C: Same as Model 4A, Except No 4th Quarter

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.1022	16.5870
WZ14	-0.4140	-11.4288
WZ26	0.2224	3.5802
WZ5	-0.0369	-1.1197
PARTIC	0.0563	1.6784
GASDRYER	0.0546	1.6242
GASCOOK	0.0441	1.2219
CKPILOT	0.1040	3.3828
POOLSPA	0.2373	3.7026
COVER	-0.2383	-3.1455
SQFTI	-0.0001	-2.9600
KIDS	0.2129	4.4311
HDSF	0.1213	53.1300
W14HDSF	-0.0448	-18.6890
A65HDSF	0.0167	7.1559
PHDSF	0.0084	3.2194
PHHDSF	-0.0039	-1.7100
CDSF	-0.0107	-6.5152
W14CDSF	0.0203	12.5758
A65CDSF	-0.0055	-3.5736
PCDSF	-0.0003	-0.1608
PHCDSF	0.0031	1.4605
Y98Q1	0.2715	8.1100
Y98Q2	-0.0010	-0.0294
Y98Q3	0.0287	0.9070
POSTQ1	-0.0631	-1.3331
POSTQ2	-0.0894	-2.1190
POSTQ3	-0.0915	-2.7217
Adjusted R-Squared:		0.5802
Number of Observations:		24725
Number of Accounts:		989
Billing Records Per Account:		25
Autoregressive Parameter		-0.73730
t-statistic		-171.51

The last model, Model 4D (Table 3-12), reverts to an earlier specification (Model 3A), but it only uses the first nine months of 1998 in the analysis. This model produces estimates of the equipment, baseline weather, and demographic variables that are similar to the previous model. The weather sensitive parameters capturing program impacts turn statistically insignificant, and most effects are captured in the binary variables.

Table 3-12
Model 4D: Same as Model 3B, Except No 4th Quarter

Parameter Estimates and Confidence Statistics		
Variable	***Coefficient***	***t-stat***
TCONST	1.0809	15.9729
WZ14	-0.4248	-11.5325
WZ26	0.2168	3.4763
WZ5	-0.0421	-1.2595
PARTIC	0.0585	1.5216
GASDRYER	0.0549	1.6143
GASCOOK	0.0433	1.1871
CKPILOT	0.1043	3.3639
POOLSPA	0.2365	3.6454
COVER	-0.2392	-3.1192
SQFTI	-0.0001	-3.1205
KIDS	0.2118	4.4170
HDSF	0.1238	54.9457
W14HDSF	-0.0457	-19.0172
A65HDSF	0.0169	7.2049
PHDSF	0.0070	2.6516
PHHDSF	-0.0022	-0.9995
CDSF	-0.0091	-6.3490
W14CDSF	0.0208	12.8489
A65CDSF	-0.0056	-3.5941
PCDSF	0.0003	0.1641
PHCDSF	-0.0010	-0.4420
Y1998	0.1114	3.6245
POSTHF98	-0.0472	-1.0848
Adjusted R-Squared:		0.5742
Number of Observations:		24725
Number of Accounts:		989
Billing Records Per Account:		25
Autoregressive Parameter		-0.73980
t-statistic		-172.81

STATISTICAL ISSUES

Correction Technique for First Order Autocorrelation

Based on the Durbin-Watson test, first order autocorrelation was found to be present in the dataset used to estimate the models specified above. For Model 1A (Table 3-2) the Durbin-Watson Statistic was .52. The critical value at a 5% significance level for the lower limit (k=21 and n=200) is 1.55, indicating an autocorrelated error structure in the data. All model specifications yielded a numerically similar statistic. Although the parameter estimates continue to be unbiased in the presence of autocorrelation, the associated standard errors will be biased. Thus, any inference concerning the statistical significance of the coefficient is unreliable. In order to correct for the statistical problem connected with autocorrelation, the standard data transformation and estimation procedure was employed. The procedure is outlined below.

The original equation was first estimated using Ordinary Least Squares (OLS) regression techniques on the time series cross sectional dataset. The residuals from this estimated equation were then computed and an average autocorrelation coefficient (ρ) was computed (SAS does this in Proc Autoreg where the p is computed using the Yule-Walker method; see SAS/ETS User's Guide, Versions 6, Second Edition, September 1996). The estimated ρ for Model 1A was -0.72. All model specifications yielded a numerically similar statistic. These values are reported with the parameter estimates. Both the dependent and independent variables (including the constant term) were then transformed by the standard transformation ($Z_1 - \rho Z_1$).

The transformed dataset, having been purged of autocorrelation, was used to re-estimate the original equation by OLS and to test for heteroskedasticity.

Testing and Correction Technique for Heteroskedasticity

Evidence of some form of heteroskedasticity was tested for using a White Test (see Greene, 1993, p. 391). Under the White Test, a test statistic is computed using the R^2 from a regression of the squared errors from an estimated equation, on all the independent variables, their respective squares, and all cross terms. The actual test statistic is computed as (total number of observations) times (the R^2 referred to above) and is distributed as a Chi-Squared variable with degrees of freedom equal to the number of independent variables.

The computed value of the relevant Chi-Squared statistic under the null hypothesis of homoskedasticity was 1077 for MODEL 1A (Table 3-2), which leads to the clear rejection of homoskedasticity in the transformed data. The high value of this test statistic was consistently reflected in all model specifications.

In the presence of heteroskedasticity, OLS yields unbiased parameter estimates. The standard errors, however, will be biased under OLS. The White Correction was used to correct the standard errors of the coefficients computed under OLS.

The corrected standard errors are unbiased and consistent. They were used to

compute the corrected t-statistics used for inferences concerning the statistical significance of the independent variables in the model. The t-statistics reported in Tables 3-2 through 3-12 were computed by dividing each coefficient estimate by its corresponding corrected standard error(square root of the coefficient's corrected variance).

Measurement Error

There was no reason to expect any abnormal measurement error in the data. Many of the variables used are binary and thus not prone to measurement error, which stems from recording errors. The square foot measure is subject to measurement error but the direction of any potential bias is rather difficult to evaluate in a multivariate regression context. Thus, there was no attempt to correct for measurement error in the analysis.

Collinearity

Generally, collinearity did not appear to be a problem in the dataset used in the analysis, with the exception of GASHEAT and GASWHTR. As mentioned previously, over 97% of all customers in the sample had gas space and water heating. Various alternative specifications were estimated with fairly consistent results concerning the important coefficient estimates. Most of the t-statistics (corrected) were significant and the estimated coefficients had their expected signs. This may be taken as evidence that collinearity did not affect the precision of the estimation procedure to any appreciable extent.

Item Non-Response and Missing Data

Most instances of missing sample data involved missing data on a variable for an entire time series. Thus attempting to replace the missing data by some regression based interpolation procedure was not feasible. Because of this, when a crucial variable's values were missing the entire time series representing a residence was deleted.

The dwelling square feet variable was the one instance where values were imputed for missing responses. There were 121 customers out of the 1055 total respondents who did not answer the questions on the square feet of their home. These were posed as, first, an open ended question (“Approximately how many square feet of heated space is there in your residence?’) and, second, as a question asking respondents to estimate the range of square feet.

To impute values for those who did not answer either question, we estimated a regression model on those who had responded to the question. The explanatory variables in the model were binary indicators of weather zone and the number of bedrooms in the house. The r-squared for this auxiliary was .31.

The estimated regression model parameters were used to impute values for the 121 cases. In the course of the regression analysis of gas consumption, we estimated the parameters of selected specifications, using the square feet with and without the observations with imputed values. We found that the exclusion of these cases had little effect on the numerical values or statistical performance of the other parameters. As a consequence, we

used the square feet variable with the imputed values throughout most of the analysis.

Treatment of Outliers

During the initial screening process of the data preparation stage of this evaluation, observations corresponding to billing data that was either implausible or extreme were omitted. A detailed discussion may be found in Chapter 2 of this report.

ESTIMATES OF NET SAVINGS

The net annual gas savings estimates based on the different models are presented in Table 3-13 along with the standard errors of the estimates and their precision at the 90% and 80% confidence levels. The savings estimates were calculated using the parameters for the variables that capture the net program impact (POSTHF98, PHHDSF, and PHCDSF), along with the values for the heating and cooling degree-days from the sample. Thus, the formula for the net program savings is:

$$\text{Savings} = \beta_1 * \text{PHHDSF} + \beta_2 * \text{PHCDSF} + \beta_3 * \text{POSTHF98}$$

Table 3-13
Estimates of Savings Attributed to the 1997 HEF Program

	Therms/Year Savings Estimates			Confidence Range	
	Method 1 [1]	***Method 2*** [2]	***Std Error***	***90%***	***80%***
Model 1A	**26.02**	25.84	15.46	0.41 - 51.27	6.02 - 45.66
Model 1B	**25.46**	25.26	15.41	0.00 - 50.60	5.50 - 45.01
Model 1C	**25.63**	25.42	15.40	0.09 - 50.75	5.68 - 45.17
Model 2A	**21.30**	21.26	15.25	0.00 - 46.34	1.71 - 40.80
Model 2B	**21.31**	21.20	15.21	0.00 - 46.22	1.70 - 40.70
Model 2C	**24.28**	24.20	16.24	0.00 - 50.91	3.39 - 45.02
Model 3A	**21.54**	21.35	14.86	0.00 - 45.79	2.30 - 40.40
Model 3B	**20.83**	20.52	14.90	0.00 - 45.04	1.42 - 39.63
Model 4A	**17.26**	17.07	9.86	0.85 - 33.29	4.43 - 29.71
Model 4B	**15.35**	16.69	8.80	2.22 - 31.17	5.41 - 27.98
Model 4C	**30.31**	28.80 [3]	6.64	17.87 - 39.73	20.28 - 37.32
Model 4D	**24.99**	25.12	14.58	1.13 - 49.11	6.43 - 43.81

1 Calculated for participants for year Oct97-Sep98, as if it were all POSTHEF
2 Calculated with means for HDSF and CDSF on regression participants' POSTHEF (10/97-11/98) bills
3 Model 4C assumes zero savings in the fourth quarter of the year

The savings were calculated using two different methods. Under the first method, the "expected" savings were computed for each participant in the sample given the square feet and weather and averaged over the sample. This method takes into account any possible correlation in the distribution square feet and weather. This would be the case, for example, if the larger homes were located in the extreme climate zones.

The second method of calculating savings simply uses the mean values of HDSF and CDSF for the participants for the period from October 1997 through November 1998. While this does not correct for the possible correlation between square feet and climate, the savings estimates using these values are very close to the savings estimates based on the other method. With only two exceptions, the estimates using the second method are also lower in magnitude than those using the first method.

The one advantage using the second method is that it allows computation of the standard errors of the estimates directly from the correct (White) covariance matrix of the parameter estimates. These standard errors are presented in Table 3-13. They show a range of values, indicating that the reliability of the estimates varies significantly from model to model. The estimates with the lowest standard errors are the last few that allowed for variations in impacts by the calendar quarter.

We calculated the 90% and 80% confidence intervals for the estimates using the standard error. These are presented in Table 3-13. The critical values for the t-statistic used to calculate the upper and lower bounds are 1.645 for the 90% confidence level and 1.282 for the 80% level.

SELECTION OF "BEST" ESTIMATE OF NET PROGRAM IMPACT

For purposes of the claim for verified program savings, it is necessary to select one model and its implied estimate of net savings as the "best" or most reliable one. Our recommendation is that Model 4C (Table 3-11) be designated as that model. This model produces an estimate of net savings of 28.8 therms. This recommendation is based on both statistical and plausibility considerations. These considerations are summarized here:

- **Overall Explanatory Performance of the Models.** All of the models are very comparable in terms of their overall power in explaining variations in natural gas consumption. While some may be superior in a strict statistical sense based on an F test of overall performance, this consideration alone is not sufficient for eliminating any one of the estimated models. The R-squared for Model 4C is .58, which is in the upper range of the values obtained for all of the models.

- **Statistical Significance of Parameter Estimates.** Almost every model had one or more parameter whose estimate was not significant from a statistical standpoint. In general, this is not sufficient to reject any one specification. However, the models that had the most reliable estimates of the parameters directly affecting the estimate of net program impacts were generally considered superior. For Model 4C, the estimates for 3 of the 5 impact variables were statistically significant at the 90% confidence level while the remaining 2 were significant at the 80% level. In contrast, the estimates for at least one of the impact parameters was statistically insignificant in the other models reported here.

- **Plausibility of Parameter Estimates.** Some of the first models that were estimated were eliminated based on the plausibility of key parameter estimates. For example, the

coefficient of the variable representing gas cooking is negative and statistically significant in Model 3A.
(Table 3-7). This is not consistent with the physical relation between the ownership of gas equipment and average gas use.

- **Plausibility of Implied Impact Estimates.** The implied point estimates of net program savings range from 16.7 to 28.8 therms per year. The 1994 HEF Program Evaluation (Study ID No. 708) found an estimate of 44 therms per year in net program savings. The evaluation of the 1990 HEF Program obtained an estimate of net savings of 39 therms[3] for water and space heating energy use. This represented an 8% reduction for the participant sample. Based on these considerations, we concluded that the estimate of savings at the high end of the range is the most plausible. Model 4C produces the highest point estimate of net savings among all of the models.

[3] This is the impact estimate for the on-site portion of the program.

Bibliography

____, 1998, Protocols and Procedures for the Verification of Cost, Benefits, and Shareholder Earnings from Demand-Side Management Programs, California Public Utilities Commission Decision 93-05-063, Revised March 1998.

____, 1993, SAS/ETS user's Guide, Version 6, Second Edition, Cary, NC: SAS Institute Inc

Greene, William H., 1993, Econometric Analysis. New York: Macmillan Publishing Company.

Appendix A

Copy of HEF Program Offer Letter and Typical Report

The Gas Company®

Southern California
Gas Company
Home Energy
Fitness Program

P.O. Box 513249
ML 12D2
Los Angeles, CA
90051-1249

Dear :

Here are the results from your Home Energy Fitness Survey, giving you a graphic picture of how your natural gas usage shapes up, along with your personalized energy-saving tips. We've designed this information to help you use energy more efficiently and trim your monthly gas bills.

Also enclosed is your Home Energy Fitness Handbook, giving you the "how-to's" on most energy improvements. Together with your personalized report, the handbook will help you decide which recommendations are best for you -- and help you implement them.

Thank you for participating in the Home Energy Fitness Program.

Sincerely,

The Gas Company

P.S. If you have any questions about your Home Energy Fitness Report, please call us at 1-800-427-2200 and select 'other marketing products and programs' from the menu options.

Home Energy Fitness Report

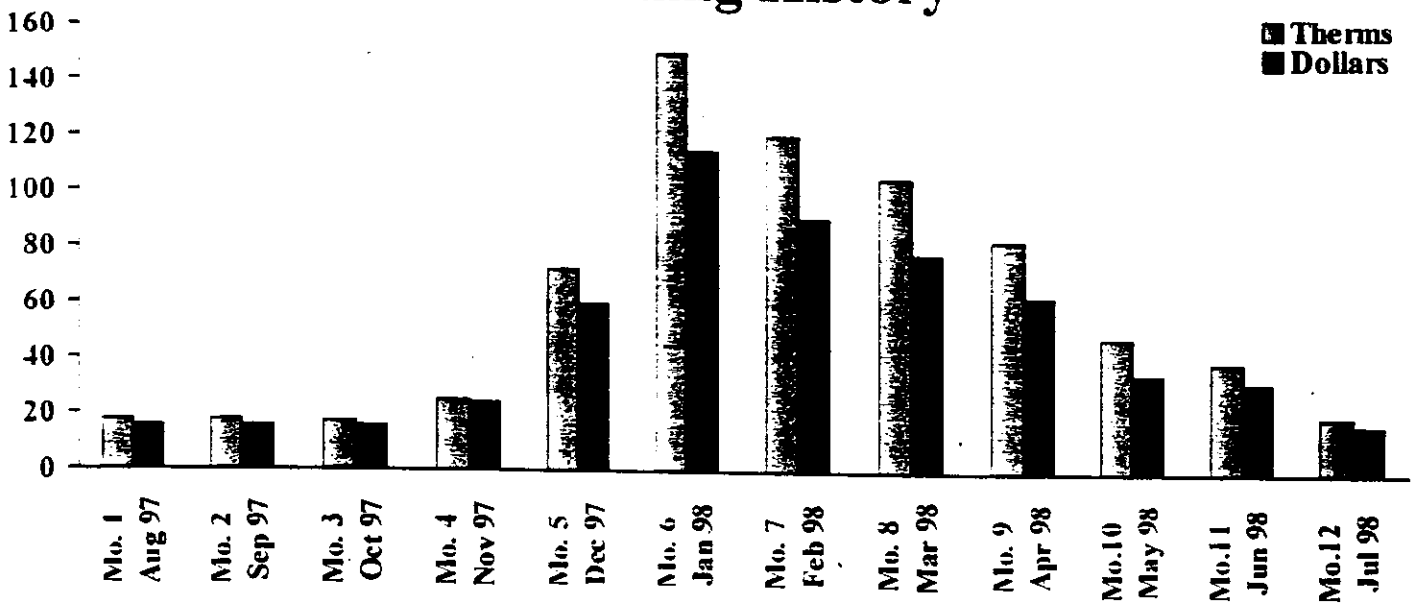
The Gas Company

Glad to be in service.

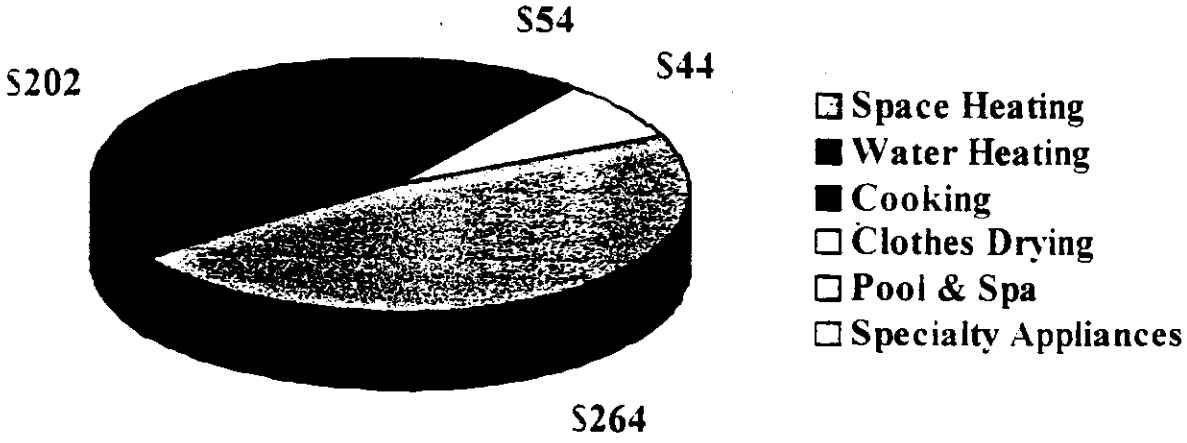
Prepared For:
Account Number:

Period Covered: Aug 1997 - Jul 1998
Total Cost: $563.60
Total Therms: 716

Gas Billing History



Gas Usage Analysis



All figures in your report are estimates only, based upon certain assumptions. Your actual figures may vary depending upon your gas usage.

The Gas Company

Glad to be of service.

Home Energy Fitness Report

Prepared For:
Account Number:

Energy Saving Recommendations

Range & Oven:

By replacing your range with a new pilotless gas range you can reduce the gas needed for cooking up to 40%. This also keeps your kitchen cooler and prevents pilot light outages caused by drafts or other disturbances.

Energy-Saving Tip: By keeping lids on your pots while cooking you can reduce your energy needs.

Clothes Dryer:

Energy-Saving Tips: 1) Follow manufacturers' loading instructions. Don't overload the dryer (which reduces energy efficiency) or run very small loads (which wastes energy) and you will save money; 2) Separate lightweight and heavy clothes for more energy-efficient drying; 3) Clean the lint trap regularly to optimize energy efficiency.

Specialty Appliances:

LOG LIGHTERS: Using a natural gas log lighter is faster and more convenient for starting log fires.

GAS LOGS: Natural gas fireplace logs offer old-fashioned charm, a safe clean-burning fuel, convenience, and cost savings. A recent national survey indicated savings of approximately 57% with gas logs versus wood logs. By switching to natural gas logs you will both reduce pollution in Southern California and save money.

GAS BBQs: A natural gas barbecue is the most economical fuel you can use and is a healthier choice for the environment. Your average energy cost per cookout with natural gas is $.07, compared to $.16 for propane and $1.68 with charcoal.

OUTDOOR HEATERS: Create a circle of warmth up to 10 feet away by installing outdoor gas heaters. The low-cost fuel allows frequent use while being easy on the environment. Outdoor gas heaters emit fewer pollutants into the atmosphere than other heating sources.

OUTDOOR LIGHTS: Outdoor gas lights are an attractive addition to the appearance of your home. The warm glow of a gas light never attracts insects while providing the equivalent of a 100 watt electric bulb that lasts for years without replacement.

The Gas Company

Glad to be of service.

Home Energy Fitness Report

Prepared For:
Account Number:

Energy Saving Recommendations

Space Heating:

Congratulations, you are saving energy and money by insulating your home. You are saving an average of 23% of your heating costs through insulation.

Energy-Saving Tips: Keep your heating system tuned up, just like you would a car. 1) For peak operating efficiency, clean and change furnace filters regularly. 2) Turn your furnace off when no one is home. 3) Loose-fitting windows and doors will lose hot or cool air through cracks around the edges. Sealing these edges with caulk and weatherstriping will keep heat and cold air where they belong. These easy practices will save $$$ and energy.

Water Heating:

Congratulations, you are saving energy and approximately 5-7% of your water heating costs by installing a high efficiency water heater(s).

Congratulations, you are saving energy and 3-5% of your water heating costs by wrapping your water heater(s).

Energy-Saving Tip: You can save even more $$$ and energy by making sure your water heater thermostat is set to 120 F.

Shower Heads:

Congratulations, you are saving energy and an average of 8-10% of your water heating costs because you have installed an energy-efficient shower head(s).

Energy-Saving Tip: You can save both water and energy by making sure that you take showers that are less than five minutes long.

Clothes Washer:

Energy-Saving Tips: 1) Whenever possible, use warm or cold settings instead of hot on your washing machine. Most laundry detergents are formulated to clean just as well at warm and cold temperatures. 2) By running your clothes washer only when it has a full load you will also save $$$ and energy.

Appendix B

Residential Telephone Survey Instrument

Hello Respondent Greeting

Hello, I'm [fill INAM] calling on behalf of Southern California Gas Company to conduct a survey of the energy practices of its residential customers. All of the information we receive will be kept strictly confidential and used only to guide our energy efficiency programs.

May I speak to [fill NAME]
<1> Person who answered is [fill NAME]
<2> [fill NAME] comes to phone
<3> Person is spouse/parent/child of [fill NAME]
<5> new number for [fill NAME]
<7> No such person/possible wrong number
<x> Callback
<y> Refused
<z> Problems--language, hearing, too ill, incapable, out of town for duration of study, etc.

Study Introduction

We are calling today to get some information about the energy practices of residential customers.

Do you have a few minutes to answer some questions?

All of the information you provide will be kept strictly confidential.

<1> PROCEED with interview
<x> Callback
<y> Refused
<z> Problems--language, hearing, too ill, incapable, out of town for duration of study, etc. [goto T162]

Call Back Introduction

Hello, I'm [fill INAM] calling on behalf of the Southern California Gas Company.

I'm calling to complete the interview we started with you…

Are you ready to start?

<1> PROCEED

<x> Callback some other time
<y> Refused this time

Answer ===>

Type of Housing

```
>q00<  Which of the following best describes your home?
            (READ LIST.  RECORD ONE RESPONSE.)

        <1>  Single family house
        <2>  Duplex or two family house
        <3>  Apartment/condominium in building with 2-4 units
        <4>  Apartment/condominium in building with 5 or more units
        <5>  Mobile home/trailer/manufactured home
        <6>  Other please specify [specify]

        Answer ===>

[If q00 ne 1 goto Thank You]
```

Heating Characteristics

```
>01<       What is the primary fuel you use for heating your home?
            (If don't know, ask "is it natural gas?")

        <1>  Natural Gas
        <3>  Other        [goto 09]
        <5>  Don't know   [goto 09]

        Answer ===>

>02<    How would you describe your attic insulation?
        Would you say you have... (READ LIST)

        <1>  Poor or no insulation
        <2>  Average insulation
        <3>  Good insulation

        <5>  Don't know

        Answer ===>

>03<    What type of thermostat do you have?
        Is it a manual thermostat or a programmable type (that can
        automatically set the temperature back at night)?

        <1>  No thermostat [goto 07]
        <2>  Manual
        <3>  Programmable/set back
        <4>  Both

        <5>  Don't know

        Answer ===>
```

>04< During the winter months,
at what temperature do you typically set your thermostat
during the daytime and evening hours when you are home.

<1> Less than 60 degrees (Turned off)
<2> 60-65 degrees

<3> 66-68 degrees
<4> 69-70 degrees
<5> 71-72 degrees
<6> 73-75 degrees
<7> Over 75 degrees

<8> Doesn't apply (Don't use it)
<9> Don't know

Answer ===>

>05< During the winter months,
at what temperature do you typically set
your thermostat at night when you go to bed?

<1> Less than 56 degrees (Turned off)
<2> 56-59 degrees
<3> 60-63 degrees
<4> 64-65 degrees
<5> 66-70 degrees
<6> Over 70 degrees

<8> Doesn't apply (Don't use it)
<9> Don't know

Answer ===>

>06< How frequently do you override or change your programmable
thermostat settings during the winter months? (READ LIST)

<1> Almost never
<2> once a week
<3> 2-4 times a week
<4> more than 4 times a week (but less than every day)
<5> every day

<9> Don't know

Answer ===>

>07< How frequently do you change your furnace filter per year?

<1> Four or more times per year 4+
<2> Three times per year 3
<3> Two times per year 2
<4> Once per year 1
<5> Less than once per year

<8> Doesn't apply (Don't have a filter, No furnace, Too old)
<9> Don't know

Answer ===>

>08< Do you turn off your furnace pilot light during the summer months?

<1> Yes
<3> No

<4> Doesn't Apply (Non-working, Don't use, Too old)
<5> Don't know

Answer ===>

>8a< In the past two years have you replaced your Furnace?

<1> Yes
<3> No
<5> No Answer or Don't Know

Answer ===>

Other Appliances

>09< Do you have a gas or electric Water Heater?
(If don't know, ask "is it natural gas?")

<1> Natural Gas [goto 9a]
<3> Other
<5> Don't know

Answer ===> [goto 10]

>9a< In the past two years have you replaced your Water Heater?

<1> Yes
<3> No
<5> No Answer or Don't Know

Answer ===>

>10< Do you have a gas clothes dryer?

<1> Yes [goto 10a]
<3> No
<5> Don't know

Answer ===> [goto 11]

>10a< In the past two years have you replaced your Clothes Dryer?

<1> Yes
<3> No
<5> No Answer or Don't Know

Answer ===>

>11< What is the primary fuel you use for cooking?
(Stove, Cooktop or Oven. Not microwave)

<1> Natural Gas [goto 12]
<3> Other
<5> Don't know

Answer ===> [goto 13]

>12< Does any of your cooking equipment have continuously-operating pilot lights?

<1> Yes
<3> No
<5> No Answer or Don't Know

Answer ===>

>12a< In the past two years have you replaced any Cooking Equipment?
(Stove top or Oven)
<1> Yes [goto 12b]
<3> No
<5> Don't know

Answer ===> [goto 13]

>12b< What kind of Cooking Equipment did you replace?

<1> Stove top (Burners)
<2> Oven(s)
<3> Both

<5> No Answer or Don't Know

Answer ===>

>13< Do you have a gas-heated swimming pool?

<1> Yes
<3> No [goto 15]

Answer ===>

>14< Do you use a cover on your pool?

<1> Yes
<3> No
<5> No Answer or Don't Know

Answer ===>

>15< Do you have a gas-heated spa or Jacuzzi?

<1> Yes
<3> No [goto 30]

Answer ===>

>16< Do you use a cover on your spa or Jacuzzi?

<1> Yes
<3> No
<5> No Answer or Don't Know

>30< Has your household received an evaluation of your gas usage from Southern California Gas Company or any other entity in the past three years?

(This is an analysis of how your household uses gas, along with recommendations on measures you can take to save energy. The evaluation offered by Southern California Gas is called the Home Energy Fitness Survey. The Home Energy Fitness Survey is based on information you would have provided by filling out a questionnaire.)

<1> Yes, HEF study [goto 32]
<3> Yes, other energy audit

<5> Don't recall participating / Don't Know [goto 37]

Answer ===>

>31< Did the other survey cover gas usage or just electric usage?

<1> Natural Gas
<3> Electric only [goto 37]

<5> Don't know [goto 37]

Answer ===>

>32< Do you recall in what year you received the Home Energy Fitness Survey or other home energy audit? If you have received more than one HEF Survey/audit, in what year was the most recent one?

<95-97> Year
<99> Don't know

Answer ===>

>36< On a scale of 1 to 6, where 1 represents "not at all helpful" and 6 "very helpful", how helpful would you rate the information you received from the Survey? (HEF or other home energy audit)

Not At All Helpful					Very Helpful	Don't Know
<1>	<2>	<3>	<4>	<5>	<6>	<9>

Answer ===>

>37< Now I'm going read a list of several things you might have done that would affect your gas consumption.

Please answer Yes to these questions
ONLY if you did them IN THE LAST 2 YEARS.

(Enter <.> When Done) ===>

>40< Did you replace any standard showerheads with a Low-flow model?
(In the last two years)

<1> Yes
<3> No
<5> No Answer or Don't Know

Answer ===>

```
>41<  Did you install Water heater wrap?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>

>42<  Did you install Pipe insulation?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>

>43<  Did you install Faucet aerators?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>

>48<  Did you lower the Water heater thermostat?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know


>50<  Did you install Attic insulation?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know


        Answer ===>


>51<  Did you install Wall insulation?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>
```

```
>52<  Did you install Floor insulation?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>

>53<  Did you install Caulk/weather-strip?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>

>54<  Did you install Door sweeps?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>


>55<  Did you install Wall socket sealers?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>

>56<  Did you install Programmable thermostat(s)
      to replace a manual one?
        (In the last two years)

          <1>  Yes
          <3>  No
          <5>  No Answer or Don't Know

        Answer ===>
```

Housing characteristics

>81< Approximately how many square feet of your home is gas-heated?

<100 - 6000> square feet [goto 83]

<9> Don't Know

Answer ===>

>82< Which range is closest to the square footage of gas-heated enclosed space of your home. (READ LIST)

<1> Under 600 square feet
<2> 601 - 1,000 square feet
<3> 1,001 - 1,500 square feet
<4> 1,501 - 2,000 square feet
<5> 2,001 or more square feet

<9> Don't Know

Answer ===>

>83< How many bedrooms are in your home?

<1-5> One to Five or more

Answer ===>

>84< Approximately what year was your home built?

<1> 1980-Present
<2> 1970-1979
<3> 1960-1969
<4> 1950-1959
<5> 1930-1949
<6> Before 1930

<9> Don't know

Answer ===>

Demographics

(The last few questions about your household are for statistical purposes only. All individual responses are strictly confidential.)

>90< Including yourself, how many people live in your household for the majority of the year?

<1> One
<3> More than One [goto 91b1]

<5> No response/Refused [goto 92]

Answer ===>

>91a< Are you: (READ LIST)

<1> 65 or older
<2> 21 to 64 years old
<3> 20 or younger

<5> No response/Refused

Answer ===> [goto 92]

>91b1< Of these individuals, how many fall into each of the following age categories?

<0-10> 65 or older

Answer ===>

>91b2< <0-10> 21 to 64 years old

Answer ===>

>91b3< <0-10> 6 to 20 years old

Answer ===>

>91b4< <0-10> Less than 6 years old

Answer ===>

>92< What is the highest level of education completed by anyone living in your household? (READ LIST)

<1> Some high school
<2> High school graduate or equivalent
<3> Some college or technical school
<4> College graduate
<5> Post-graduate study
<9> No response/refused

Answer ===>

>93< Which of the following best describes your total household annual income before taxes and other deductions? (READ LIST)

<1> Under $10,000
<2> Under $20,000
<3> Under $30,000
<4> Under $40,000
<5> Under $50,000

<6> Under $75,000
<7> $75,000 or more

<9> No response/refused

Answer ===>

Thank You Ending

Thank you [fill NAME] very much for giving us your time today.

We appreciate your help with this study.

(HANG UP LINE)

Any notes for supervisor or for coders before you finish with this case?

<1> Yes
<3> No

Answer ===>

End Home Energy Fitness Program Survey Script

Appendix C

Tabulations of Telephone Survey Responses

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Label
53	ACCTNO	Num	8	416	BEST12.	Account Number
59	ADDRESS	Char	30	488		Service Address
65	AUDITDT	Num	8	551	MMDDYY8.	Participation Date
57	CITY	Char	16	452		City
55	FNAME	Char	9	429		First Name
58	LNAME	Char	20	468		Last Name
67	NPSSMPL	Num	8	567	NOYES.	In NPart Survey Sample
66	PARTIC	Num	8	559	PART.	Participant Status
56	PHONE	Char	14	438		Phone Number
1	Q00	Num	8	0	BTYPE.	Building Type
2	Q01	Num	8	8	FUEL.	Primary Heating Fuel
3	Q02	Num	8	16	QUAL.	Attic Insulation
4	Q03	Num	8	24	TTYPE.	Thermostat Type
5	Q04	Num	8	32	TMP1S.	Day Temperature, When Home
6	Q05	Num	8	40	TMP2S.	Night Temperature
7	Q06	Num	8	48	FREQO.	Override Thermostat
8	Q07	Num	8	56	FREQF.	Change Furnace Filter
9	Q08	Num	8	64	YESNO.	Turn Off Pilot, Summer?
11	Q09	Num	8	80	FUEL.	Water Heating Fuel
13	Q10	Num	8	96	YESNO.	Have Gas Clothes Dryer?
15	Q11	Num	8	112	FUEL.	Primary Cooking Fuel
16	Q12	Num	8	120	YESNO.	Pilot on Cooking Equipment?
19	Q13	Num	8	144	YESNO.	Have Gas-Heated Swimming Pool?
20	Q14	Num	8	152	YESNO.	Use Cover on Swimming Pool?
21	Q15	Num	8	160	YESNO.	Have Gas-Heated Spa/Jacuzzi?
22	Q16	Num	8	168	YESNO.	Use Cover on Spa/Jacuzzi?
23	Q30	Num	8	176	YNHEF.	Received Evaluatn Last 3 Yrs?
24	Q31	Num	8	184	GEFUL.	Other Audit: Fuel Covered
25	Q32	Num	8	192	HEFYR.	Last HEF Year, Recalled
26	Q36	Num	8	200	SCALE.	Scale: How Helpful Was Audit?
27	Q40	Num	8	208	YESNO.	Install Low-Flow Showerheads?
28	Q41	Num	8	216	YESNO.	Install Water Heater Wrap?
29	Q42	Num	8	224	YESNO.	Install Pipe Insulation?
30	Q43	Num	8	232	YESNO.	Install Faucet Aerators?
31	Q48	Num	8	240	YESNO.	Lower Water Heater Temperature?
32	Q50	Num	8	248	YESNO.	Install Attic Insulation?
33	Q51	Num	8	256	YESNO.	Install Wall Insulation?
34	Q52	Num	8	264	YESNO.	Install Floor Insulation?
35	Q53	Num	8	272	YESNO.	Install Caulk/Wtherstripping?
36	Q54	Num	8	280	YESNO.	Install Door Sweeps?
37	Q55	Num	8	288	YESNO.	Install Wall Socket Sealers?
38	Q56	Num	8	296	YESNO.	Install Setback T-Stat(s)?
39	Q81	Num	8	304	BEST12.	Gas-Heated Square Footage
40	Q82	Num	8	312	SQFT.	Gas-Heated SqFt Estimate
41	Q83	Num	8	320	ROOMS.	Number of Bedrooms
42	Q84	Num	8	328	VINT.	Year That Home Was Built
43	Q90	Num	8	336	ONEP.	# Persons in Household
49	Q92	Num	8	384	EDUC.	Highest Level Education in HH
50	Q93	Num	8	392	INCM.	Total HH Pre-Tax Annual Income
14	Q10A	Num	8	104	YESNO.	Replaced Dryer Last 2 Yrs?
17	Q12A	Num	8	128	YESNO.	Replaced Cook Eq Last 2 Yrs?
18	Q12B	Num	8	136	CKEQ.	Which Cook Equipmt Replaced
10	Q8A	Num	8	72	YESNO.	Replaced Furnace Last 2 Yrs?
44	Q91A	Num	8	344	AGEG.	Age Group of Sole Resident

2

CONTENTS PROCEDURE

#	Variable	Type	Len	Pos	Format	Label
45	Q91B1	Num	8	352	BEST12.	# Residents 65 or Older
46	Q91B2	Num	8	360	BEST12.	# Residents 21-64 Yrs Old
47	Q91B3	Num	8	368	BEST12.	# Residents 6-20 Yrs Old
48	Q91B4	Num	8	376	BEST12.	# Residents <6 Yrs Old
12	Q9A	Num	8	88	YESNO.	Replaced Wtr Htr Last 2 Yrs?
64	QSTRATUM	Char	1	550		Strata for Quota (1-23)
52	SID	Num	8	408	BEST12.	Survey ID
51	SPANISH	Num	8	400	NOYES.	Spanish speaking
63	STRATUM	Num	8	542	BEST12.	Strata Number (1-36)
61	TOYCAT	Num	8	526	TOYCAT.	Turn-On Year Strata
62	UPDCAT	Num	8	534	UPDCAT.	Use Per Day Strata
60	WZONE	Num	8	518	WZONE.	Weather Zone
54	ZIPCODE	Char	5	424		Zip Code

3

TABLE OF PARTIC BY NPSSMPL

PARTIC(Participant Status)
NPSSMPL(In NPart Survey Sample)

Frequency / Row Pct	No	Yes	Total
NonPart	0 0.00	503 100.00	503
Partic	549 96.32	21 3.68	570
Total	549	524	1073

TABLE OF PARTIC BY Q00

PARTIC(Participant Status)
Q00(Building Type)

Frequency / Row Pct	single-family	Total
NonPart	503 100.00	503
Partic	570 100.00	570
Total	1073	1073

TABLE OF PARTIC BY Q01

PARTIC(Participant Status) Q01(Primary Heating Fuel)

Frequency / Row Pct	natural gas	other	dk	Total
NonPart	490 97.42	11 2.19	2 0.40	503
Partic	563 98.77	7 1.23	0 0.00	570
Total	1053	18	2	1073

TABLE OF PARTIC BY Q02

PARTIC(Participant Status) Q02(Attic Insulation)

Frequency / Row Pct	poor	average	good	dk	Total
NonPart	51 10.41	162 33.06	248 50.61	29 5.92	490
Partic	45 7.99	193 34.28	298 52.93	27 4.80	563
Total	96	355	546	56	1053

Frequency Missing = 20

4

TABLE OF PARTIC BY Q03

PARTIC(Participant Status) Q03(Thermostat Type)

Frequency / Row Pct	none	manual	setback	both	dk	Total
NonPart	15	306	150	16	3	490
	3.06	62.45	30.61	3.27	0.61	
Partic	10	335	194	21	3	563
	1.78	59.50	34.46	3.73	0.53	
Total	25	641	344	37	6	1053

Frequency Missing = 20

TABLE OF PARTIC BY Q04

PARTIC(Participant Status) Q04(Day Temperature, When Home)

Frequency / Row Pct	<60F	60-65F	66-68F	69-70F	71-72F	73-75F	>75F	n/a	dk	Total
NonPart	38	51	101	163	75	30	7	6	4	475
	8.00	10.74	21.26	34.32	15.79	6.32	1.47	1.26	0.84	
Partic	34	66	136	180	69	43	15	1	9	553
	6.15	11.93	24.59	32.55	12.48	7.78	2.71	0.18	1.63	
Total	72	117	237	343	144	73	22	7	13	1028

Frequency Missing = 45

TABLE OF PARTIC BY Q05

PARTIC(Participant Status) Q05(Night Temperature)

Frequency / Row Pct	<56F	56-59F	60-63F	64-65F	66-70F	>70F	n/a	dk	Total
NonPart	202	35	66	59	71	28	6	8	475
	42.53	7.37	13.89	12.42	14.95	5.89	1.26	1.68	
Partic	222	28	97	65	100	32	1	8	553
	40.14	5.06	17.54	11.75	18.08	5.79	0.18	1.45	
Total	424	63	163	124	171	60	7	16	1028

Frequency Missing = 45

The SAS System
5

TABLE OF PARTIC BY Q06

PARTIC(Participant Status) Q06(Override Thermostat)

Frequency Row Pct	almost n ever	once a w eek	2-4 time s/wk	>4 times /wk	every da y	dk	Total
NonPart	116 69.88	16 9.64	16 9.64	6 3.61	6 3.61	6 3.61	166
Partic	151 70.23	21 9.77	19 8.84	7 3.26	13 6.05	4 1.86	215
Total	267	37	35	13	19	10	381

Frequency Missing = 692

TABLE OF PARTIC BY Q07

PARTIC(Participant Status) Q07(Change Furnace Filter)

Frequency Row Pct	4+ times /yr	3 times/ yr	2 times/ yr	once a y ear	<once a yr	n/a	dk	Total
NonPart	145 29.59	34 6.94	72 14.69	114 23.27	23 4.69	84 17.14	18 3.67	490
Partic	171 30.37	32 5.68	101 17.94	121 21.49	18 3.20	96 17.05	24 4.26	563
Total	316	66	173	235	41	180	42	1053

Frequency Missing = 20

TABLE OF PARTIC BY Q08

PARTIC(Participant Status) Q08(Turn Off Pilot, Summer?)

Frequency Row Pct	yes	no	n/a	dk	Total
NonPart	146 29.80	259 52.86	72 14.69	13 2.65	490
Partic	164 29.13	295 52.40	94 16.70	10 1.78	563
Total	310	554	166	23	1053

Frequency Missing = 20

6

TABLE OF PARTIC BY Q8A

PARTIC(Participant Status)
Q8A(Replaced Furnace Last 2 Yrs?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	48 9.80	440 89.80	2 0.41	490
Partic	66 11.72	494 87.74	3 0.53	563
Total	114	934	5	1053

Frequency Missing = 20

TABLE OF PARTIC BY Q09

PARTIC(Participant Status) Q09(Water Heating Fuel)

Frequency / Row Pct	natural gas	other	dk	Total
NonPart	491 97.61	10 1.99	2 0.40	503
Partic	554 97.19	9 1.58	7 1.23	570
Total	1045	19	9	1073

TABLE OF PARTIC BY Q9A

PARTIC(Participant Status)
Q9A(Replaced Wtr Htr Last 2 Yrs?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	85 17.31	404 82.28	2 0.41	491
Partic	125 22.56	429 77.44	0 0.00	554
Total	210	833	2	1045

Frequency Missing = 28

TABLE OF PARTIC BY Q10

PARTIC(Participant Status) Q10(Have Gas Clothes Dryer?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	381 75.75	118 23.46	4 0.80	503
Partic	448 78.60	119 20.88	3 0.53	570
Total	829	237	7	1073

The SAS System
7

TABLE OF PARTIC BY Q10A

PARTIC(Participant Status)
Q10A(Replaced Dryer Last 2 Yrs?)

Frequency³ Row Pct ³	yes	no	dk	Total
NonPart ³	66	314	1	381
³	17.32	82.41	0.26	
Partic ³	79	368	1	448
³	17.63	82.14	0.22	
Total	145	682	2	829

Frequency Missing = 244

TABLE OF PARTIC BY Q11

PARTIC(Participant Status)
Q11(Primary Cooking Fuel)

Frequency³ Row Pct ³	natural gas ³	other ³	Total
NonPart ³	423	80	503
³	84.10	15.90	
Partic ³	463	107	570
³	81.23	18.77	
Total	886	187	1073

TABLE OF PARTIC BY Q12

PARTIC(Participant Status)
Q12(Pilot on Cooking Equipment?)

Frequency³ Row Pct ³	yes	no	dk	Total
NonPart ³	130	292	1	423
³	30.73	69.03	0.24	
Partic ³	121	339	3	463
³	26.13	73.22	0.65	
Total	251	631	4	886

Frequency Missing = 187

The SAS System
8

TABLE OF PARTIC BY Q12A

PARTIC(Participant Status)
Q12A(Replaced Cook Eq Last 2 Yrs?)

Frequency³ Row Pct ³	yes	no	dk	Total
NonPart ³	59	364	0	423
³	13.95	86.05	0.00	
Partic ³	62	400	1	463
³	13.39	86.39	0.22	
Total	121	764	1	886

Frequency Missing = 187

TABLE OF PARTIC BY Q12B

PARTIC(Participant Status) Q12B(Which Cook Equipmt Replaced)

Frequency³ Row Pct ³	stove to p	oven	both	na/dk	Total
NonPart ³	36	4	19	0	59
³	61.02	6.78	32.20	0.00	
Partic ³	42	9	9	2	62
³	67.74	14.52	14.52	3.23	
Total	78	13	28	2	121

Frequency Missing = 952

TABLE OF PARTIC BY Q13

PARTIC(Participant Status)
Q13(Have Gas-Heated Swimming Pool?)

Frequency³ Row Pct ³	yes	no	Total
NonPart ³	43	460	503
³	8.55	91.45	
Partic ³	68	502	570
³	11.93	88.07	
Total	111	962	1073

9

TABLE OF PARTIC BY Q14

PARTIC(Participant Status)
Q14(Use Cover on Swimming Pool?)

Frequency / Row Pct	yes	no	Total
NonPart	20 46.51	23 53.49	43
Partic	33 48.53	35 51.47	68
Total	53	58	111

Frequency Missing = 962

TABLE OF PARTIC BY Q15

PARTIC(Participant Status)
Q15(Have Gas-Heated Spa/Jacuzzi?)

Frequency / Row Pct	yes	no	Total
NonPart	82 16.30	421 83.70	503
Partic	103 18.07	467 81.93	570
Total	185	888	1073

TABLE OF PARTIC BY Q16

PARTIC(Participant Status)
Q16(Use Cover on Spa/Jacuzzi?)

Frequency / Row Pct	yes	no	Total
NonPart	54 65.85	28 34.15	82
Partic	68 66.02	35 33.98	103
Total	122	63	185

Frequency Missing = 888

TABLE OF PARTIC BY Q30

PARTIC(Participant Status)
Q30(Received Evaluatn Last 3 Yrs?)

Frequency / Row Pct	yes, HEF study	yes, oth er audit	don't re call/dk	Total
NonPart	43 8.55	3 0.60	457 90.85	503
Partic	92 16.14	11 1.93	467 81.93	570
Total	135	14	924	1073

The SAS System
10

TABLE OF PARTIC BY Q31

PARTIC(Participant Status)
Q31(Other Audit: Fuel Covered)

Frequency / Row Pct	natural gas	elec only	dk	Total
NonPart	1 33.33	2 66.67	0 0.00	3
Partic	1 9.09	8 72.73	2 18.18	11
Total	2	10	2	14

Frequency Missing = 1059

TABLE OF PARTIC BY Q32

PARTIC(Participant Status) Q32(Last HEF Year, Recalled)

Frequency / Row Pct	1995	1996	1997	dk	Total
NonPart	5 11.36	16 36.36	17 38.64	6 13.64	44
Partic	11 11.83	32 34.41	43 46.24	7 7.53	93
Total	16	48	60	13	137

Frequency Missing = 936

TABLE OF PARTIC BY Q36

PARTIC(Participant Status) Q36(Scale: How Helpful Was Audit?)

Frequency / Row Pct	1:not	2	3	4	5	6:very	dk	Total
NonPart	2 4.55	1 2.27	5 11.36	5 11.36	10 22.73	16 36.36	5 11.36	44
Partic	4 4.30	4 4.30	15 16.13	12 12.90	13 13.98	41 44.09	4 4.30	93
Total	6	5	20	17	23	57	9	137

Frequency Missing = 936

The SAS System
11

TABLE OF PARTIC BY Q40

PARTIC(Participant Status)
Q40(Install Low-Flow Showerheads?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	208 41.35	292 58.05	3 0.60	503
Partic	246 43.16	321 56.32	3 0.53	570
Total	454	613	6	1073

TABLE OF PARTIC BY Q41

PARTIC(Participant Status)
Q41(Install Water Heater Wrap?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	130 25.84	369 73.36	4 0.80	503
Partic	151 26.49	417 73.16	2 0.35	570
Total	281	786	6	1073

TABLE OF PARTIC BY Q42

PARTIC(Participant Status)
Q42(Install Pipe Insulation?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	53 10.54	446 88.67	4 0.80	503
Partic	66 11.58	499 87.54	5 0.88	570
Total	119	945	9	1073

TABLE OF PARTIC BY Q43

PARTIC(Participant Status)
Q43(Install Faucet Aerators?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	138 27.44	362 71.97	3 0.60	503
Partic	150 26.32	417 73.16	3 0.53	570
Total	288	779	6	1073

The SAS System
12
TABLE OF PARTIC BY Q48

PARTIC(Participant Status)
Q48(Lower Water Heater Temperature?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	184 36.58	302 60.04	17 3.38	503
Partic	193 33.86	358 62.81	19 3.33	570
Total	377	660	36	1073

TABLE OF PARTIC BY Q50

PARTIC(Participant Status)
Q50(Install Attic Insulation?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	66 13.12	435 86.48	2 0.40	503
Partic	71 12.46	499 87.54	0 0.00	570
Total	137	934	2	1073

TABLE OF PARTIC BY Q51

PARTIC(Participant Status)
Q51(Install Wall Insulation?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	30 5.96	472 93.84	1 0.20	503
Partic	41 7.19	528 92.63	1 0.18	570
Total	71	1000	2	1073

TABLE OF PARTIC BY Q52

PARTIC(Participant Status)
Q52(Install Floor Insulation?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	26 5.17	474 94.23	3 0.60	503
Partic	20 3.51	550 96.49	0 0.00	570
Total	46	1024	3	1073

The SAS System
13

TABLE OF PARTIC BY Q53

PARTIC(Participant Status)
Q53(Install Caulk/Wtherstripping?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	129	371	3	503
	25.65	73.76	0.60	
Partic	155	413	2	570
	27.19	72.46	0.35	
Total	284	784	5	1073

TABLE OF PARTIC BY Q54

PARTIC(Participant Status) Q54(Install Door Sweeps?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	115	387	1	503
	22.86	76.94	0.20	
Partic	116	453	1	570
	20.35	79.47	0.18	
Total	231	840	2	1073

TABLE OF PARTIC BY Q55

PARTIC(Participant Status)
Q55(Install Wall Socket Sealers?)

Frequency / Row Pct	yes	no	dk	Total
NonPart	27	475	1	503
	5.37	94.43	0.20	
Partic	31	534	5	570
	5.44	93.68	0.88	
Total	58	1009	6	1073

TABLE OF PARTIC BY Q56

PARTIC(Participant Status)
Q56(Install Setback T-Stat(s)?)

Frequency / Row Pct	yes	no	Total
NonPart	50	173	223
	22.42	77.58	
Partic	60	212	272
	22.06	77.94	
Total	110	385	495

Frequency Missing = 578

The SAS System
16

TABLE OF PARTIC BY Q82

PARTIC(Participant Status) Q82(Gas-Heated SqFt Estimate)

Frequency / Row Pct	<600 sf	601-1000	1001-150	1501-200	>2000 sf	dk	Total
NonPart	2	6	6	2	1	62	79
	2.53	7.59	7.59	2.53	1.27	78.48	
Partic	1	2	7	4	1	60	75
	1.33	2.67	9.33	5.33	1.33	80.00	
Total	3	8	13	6	2	122	154

Frequency Missing = 919

TABLE OF PARTIC BY Q83

PARTIC(Participant Status) Q83(Number of Bedrooms)

Frequency / Row Pct	1	2	3	4	5+	Total
NonPart	2	68	274	142	17	503
	0.40	13.52	54.47	28.23	3.38	
Partic	2	73	320	142	33	570
	0.35	12.81	56.14	24.91	5.79	
Total	4	141	594	284	50	1073

TABLE OF PARTIC BY Q84

PARTIC(Participant Status) Q84(Year That Home Was Built)

Frequency / Row Pct	1980-pre sent	1970-79	1960-69	1950-59	1930-49	before 1930	dk	Total
NonPart	93	83	91	103	60	22	51	503
	18.49	16.50	18.09	20.48	11.93	4.37	10.14	
Partic	124	111	105	119	42	23	46	570
	21.75	19.47	18.42	20.88	7.37	4.04	8.07	
Total	217	194	196	222	102	45	97	1073

TABLE OF PARTIC BY Q90

PARTIC(Participant Status) Q90(# Persons in Household)

Frequency / Row Pct	one	two+	nr	Total
NonPart	48	432	23	503
	9.54	85.88	4.57	
Partic	62	480	28	570
	10.88	84.21	4.91	
Total	110	912	51	1073

The SAS System
17

TABLE OF PARTIC BY Q91A

PARTIC(Participant Status)
Q91A(Age Group of Sole Resident)

Frequency / Row Pct	65+ yrs old	21-64 yrs	Total
NonPart	32 66.67	16 33.33	48
Partic	51 82.26	11 17.74	62
Total	83	27	110

Frequency Missing = 963

TABLE OF PARTIC BY Q91B1

PARTIC(Participant Status) Q91B1(# Residents 65 or Older)

Frequency / Row Pct	0	1	2	3	9	Total
NonPart	272 62.96	53 12.27	105 24.31	2 0.46	0 0.00	432
Partic	273 56.88	60 12.50	144 30.00	2 0.42	1 0.21	480
Total	545	113	249	4	1	912

Frequency Missing = 161

TABLE OF PARTIC BY Q91B2

PARTIC(Participant Status) Q91B2(# Residents 21-64 Yrs Old)

Frequency / Row Pct	0	1	2	3	4	5	6	7	10	Total
NonPart	75 17.36	50 11.57	218 50.46	73 16.90	14 3.24	0 0.00	1 0.23	1 0.23	0 0.00	432
Partic	118 24.58	66 13.75	223 46.46	57 11.88	9 1.88	3 0.63	2 0.42	1 0.21	1 0.21	480
Total	193	116	441	130	23	3	3	2	1	912

Frequency Missing = 161

The SAS System
18

TABLE OF PARTIC BY Q91B3

PARTIC(Participant Status) Q91B3(# Residents 6-20 Yrs Old)

Frequency / Row Pct	0	1	2	3	4	5	6	Total
NonPart	248 57.41	86 19.91	69 15.97	21 4.86	4 0.93	4 0.93	0 0.00	432
Partic	306 63.75	81 16.88	67 13.96	18 3.75	7 1.46	0 0.00	1 0.21	480
Total	554	167	136	39	11	4	1	912

Frequency Missing = 161

TABLE OF PARTIC BY Q91B4

PARTIC(Participant Status) Q91B4(# Residents <6 Yrs Old)

Frequency / Row Pct	0	1	2	3	Total
NonPart	386 89.35	40 9.26	6 1.39	0 0.00	432
Partic	438 91.25	36 7.50	4 0.83	2 0.42	480
Total	824	76	10	2	912

Frequency Missing = 161

TABLE OF PARTIC BY Q92

PARTIC(Participant Status) Q92(Highest Level Education in HH)

Frequency / Row Pct	some hig h school	high sch ool grad	some col lege	college grad	post-gra d study	nr	Total
NonPart	18 3.58	77 15.31	136 27.04	175 34.79	75 14.91	22 4.37	503
Partic	14 2.46	82 14.39	150 26.32	195 34.21	93 16.32	36 6.32	570
Total	32	159	286	370	168	58	1073

The SAS System
19

TABLE OF PARTIC BY Q93

PARTIC(Participant Status) Q93(Total HH Pre-Tax Annual Income)

Frequency / Row Pct	<$10K/yr	$10K-$20K	$20K-$30K	$30K-$40K	$40K-$50K	$50K-$75K	>$75K/yr	nr	Total
NonPart	11	30	37	37	55	74	65	194	503
	2.19	5.96	7.36	7.36	10.93	14.71	12.92	38.57	
Partic	9	21	32	50	59	82	68	249	570
	1.58	3.68	5.61	8.77	10.35	14.39	11.93	43.68	
Total	20	51	69	87	114	156	133	443	1073

TABLE OF PARTIC BY TOYCAT

PARTIC(Participant Status)
TOYCAT(Turn-On Year Strata)

Frequency / Row Pct	pre-1980	1980+	Total
NonPart	200	303	503
	39.76	60.24	
Partic	246	324	570
	43.16	56.84	
Total	446	627	1073

TABLE OF PARTIC BY WZONE

PARTIC(Participant Status) WZONE(Weather Zone)

Frequency / Row Pct	Mountain	Lower De sert	Coastal Strip	Upper De sert	Inland V alley	LA Basin	Total
NonPart	3	8	207	88	185	12	503
	0.60	1.59	41.15	17.50	36.78	2.39	
Partic	4	15	251	87	199	14	570
	0.70	2.63	44.04	15.26	34.91	2.46	
Total	7	23	458	175	384	26	1073

TABLE OF PARTIC BY UPDCAT

PARTIC(Participant Status) UPDCAT(Use Per Day Strata)

Frequency / Row Pct	< 1.6 th erms/day	1.6-2.0 thms/day	>= 2.0 t hrms/day	Total
NonPart	172	175	156	503
	34.19	34.79	31.01	
Partic	196	214	160	570
	34.39	37.54	28.07	
Total	368	389	316	1073

The SAS System
20

Analysis Variable : Q81 Gas-Heated Square Footage

PARTIC	N Obs	N	Mean	Std Dev	Minimum	Maximum
NonPart	503	424	1840.07	636.5081215	100.0000000	5000.00
Partic	570	495	1892.37	685.5956320	200.0000000	5400.00

M&E PROTOCOLS TABLE 6
RESULTS USED TO SUPPORT PY97 SECOND EARNINGS CLAIM

FOR

RESIDENTIAL ENERGY MANAGEMENT SERVICES PROGRAM
FIRST YEAR LOAD IMPACT EVALUATION

MARCH 1999

STUDY ID NO. 715

SOUTHERN CALIFORNIA GAS COMPANY
M&E PROTOCOLS TABLE 6 - RESULTS USED TO SUPPORT PY97 SECOND EARNINGS CLAIM FOR RESIDENTIAL ENERGY MANAGEMENT SERVICES PROGRAM
FIRST YEAR LOAD IMPACT EVALUATION, MARCH 1999, STUDY ID NO. 715

Designated Unit of Measurement: LOAD IMPACTS (THERMS/YEAR) PER PARTICIPANT
END USE: ALL GAS END USES AND PRACTICES COMBINED

				5. A. 90% CONFIDENCE LEVEL				5. B. 90% CONFIDENCE LEVEL			
				LOWER BOUND	UPPER BOUND	LOWER BOUND	UPPER BOUND	LOWER BOUND	UPPER BOUND	LOWER BOUND	UPPER BOUND
A. Pre-install usage:	Pre-Install Therms	653.0	631.0	634.4	671.6	612.4	649.6	638.5	667.5	616.5	645.5
	Base Therms	653.0	631.0	634.4	634.4	612.4	649.6	638.5	667.5	616.5	645.5
	Base Therms/ designated unit of measurement	653.0	631.0	634.4	671.6	612.4	649.6	638.5	667.5	616.5	645.5
B. Impact year usage:	Impact Yr Therms	809.0	780	790.4	827.6	761.4	798.6	794.5	823.5	765.5	794.5
	Impact Yr Therms/designated unit	809.0	780	790.4	827.6	761.4	798.6	794.5	823.5	765.5	794.5
A. Load impacts:	Load Impacts - Therms	N/A	28.8	N/A	N/A	17.9	39.7	N/A	N/A	20.3	37.3
B. Load impacts per DUOM	Load Impacts/designated unit - Therms	N/A	28.8	N/A	N/A	17.9	39.7	N/A	N/A	20.3	37.3
C. % Change	i. % change in usage - Part Grp - Therms	N/A	-4.56%	N/A	N/A	-2.84%	-6.29%	N/A	N/A	-3.22%	-5.91%
	ii. % change in usage - Comp Grp - Therms	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
D. Realization Rate:	i. Load Impacts - Therms, realization rate	N/A	65%	N/A	N/A	41%	90%	N/A	N/A	46%	85%
	ii. Load Impacts/designated unit - Therms, realization rate	N/A	65%	N/A	N/A	41%	90%	N/A	N/A	46%	85%
		RATIO		RATIO	RATIO			RATIO	RATIO		
	A. iii. Average Load Impacts - Therms	N/A		N/A	N/A			N/A	N/A		
	B. iii. Avg Load Impacts/designated unit of measurement - Therms	N/A		N/A	N/A			N/A	N/A		
	C. iii. Avg Load Impacts based on % chg in usage in Impact year relative to Base usage in Impact year - Thms	N/A		N/A	N/A			N/A	N/A		
	A. Pre-install average values			N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Gas Participant Square Footage	N/A	N/A								
	Gas Participant Number in Household	N/A	N/A								
	B. Post-install average values	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Gas Participant Square Footage	N/A	N/A								
	Gas Participant Number in Household	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	A. Number of measures installed by participants in Part Group	N/A									
	B. Number of measures installed by all program participants in the 12 months of the program year	N/A									
	C. Number of measures installed by Comp Group	N/A									
	Number of Participants	18,446	5,171								

Note1: No gross impacts were estimated because it was impossible to separate conservation measures and actions from underlying trend in gas consumption.
Note 2: Ex ante estimated savings and participant count used in realization rates are taken from the Advice Letter 2526, October 1, 1996.

M&E PROTOCOLS TABLE 7
DATA QUALITY AND PROCESSING DOCUMENTATION

FOR

RESIDENTIAL ENERGY MANAGEMENT SERVICES PROGRAM
FIRST YEAR LOAD IMPACT EVALUATION

MARCH 1999

STUDY ID NO. 715

M&E PROTOCOLS TABLE 7
DATA QUALITY AND PROCESSING DOCUMENTATION

Residential Energy Management Services Program
First Year Load Impact Evaluation
March 1999
Study ID No. 715

1. OVERVIEW INFORMATION

a. **Study Title and Study ID:** 1997 Residential Energy Management Services Program: First Year Load Impact Evaluation (Home Energy Fitness Program), Study ID No. 715, March 1999.

b. **Program, Program Year(s), and Program Description (Design):** Residential Energy Management Services Program for the 1997 program year. The Home Energy Fitness Program provides customers with comprehensive information about energy management measures and practices to reduce gas consumption.

c. **End Uses and/or Measures Covered:** All end uses combined per protocol Table C-11.

d. **Methods and Models Used:** The study uses a regression-based billing analysis to estimate net program impacts. See the sections of the report entitled “Methodology” and “Regression Model Specification” on pages 3-1 and 3-4, respectively, for a complete description of the model specifications.

e. **Participant and Comparison Group Definition:** For the load impact analysis, the participants are defined as customers who received a HEF Program report during 1997. The comparison group was drawn at random from SoCalGas residential accounts in single-family homes that had not previously participated in the HEF Program.

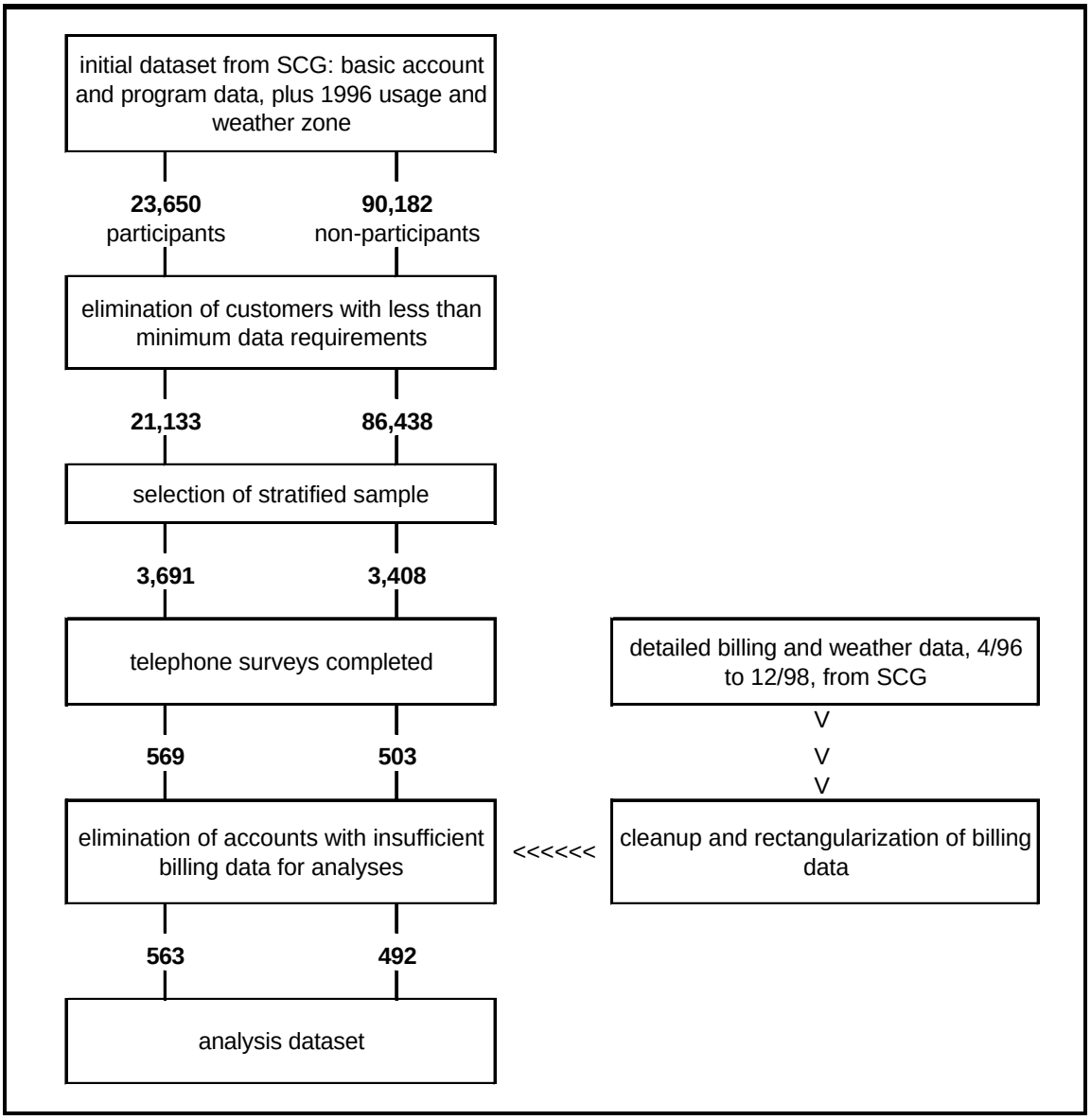
f. **Analysis Sample Size:** 563 participants and 492 nonparticipants, 27 months of billing data per customer.

2. DATABASE MANAGEMENT

a. **Data Preparation:** Data preparation, attrition, sampling and merging are documented in detail in Chapter 2 of the evaluation report.

Flow Chart:

Screening of Data for SCG 1997 HEF Evaluation Analysis



b. **Data Sources:** the data came from the following sources:

2) Participant name, address, account number, and participation date from the 1997 Home Energy Fitness Program tracking database;

3) Nonparticipant name, address, and account number from the SoCalGas billing system;

4) 1996-1998 gas consumption history from the SoCalGas billing system;

5) 1996-1998 daily weather data for various locations from newspaper reports of daily highs and lows; and

6) Participant and nonparticipant telephone surveys.

The data were merged together to form the dataset for the regression analysis leading to the estimated energy savings per dwelling unit. The data preparation process is documented in Chapter 2 of the evaluation report.

c. **Data Attrition:**

1) Participant Sample – Gas Load Impact Analysis

Number of Participants for Gas Load Impact Analysis	
1997 HEF Participants Initial Database	23,650
Remaining accounts after initial screening (residential, single family account that received audit in 1997, that was active throughout 1996)	21,133
Participants who completed telephone survey	569
Participants after final screening (complete, clean billing data for 27 months)	563

2) Nonparticipant Sample – Gas Load Impact Analysis

Number of Nonparticipants for Load Impact Analysis	
Nonparticipant initial database	90,182
Remaining accounts after initial screening (residential, single family account that had not previously received audit and was active throughout 1996)	86,438
Participants meeting minimum pre/post data requirements	2,570
Nonparticipants who completed telephone survey	503
Nonparticipants after final screening (complete, clean billing data for 27 months)	492

d. **Data Quality Checks:** The data sets for the regression analysis were merged in SAS by the appropriate key variables. Counts of the data sets before and after the merges were verified to ensure accurate merging.

e. **Data Utilization.** The completed surveys and corresponding billing and weather data were utilized for this analysis.

3. SAMPLING

a. **Sampling Procedures and Protocols:** The participant sample was a stratified, proportional random sample. The nonparticipant sample was stratified and drawn to match the distribution of the participant sample. See Chapter 2 of the evaluation report for a complete discussion and presentation of sample frequencies.

b. **Survey Information:** The survey instrument is presented in Appendix B of the evaluation report. The response rates to each question are presented in Appendix C. The disposition of the sample provided to the telephone survey implementation subcontractor are shown in Table 2-4. No tests were performed to examine possible non-response bias.

c. **Statistical Descriptions:** The mean values of all variables used in the regression model specifications are presented in Table 3-1 of the evaluation report.

4. DATA SCREENING AND ANALYSIS

a. **Outliers.** During the screening of the gas billing data, cases with extremely high (> 6500 therms) or low annual (< 200 therms) consumption were eliminated. (See Chapter 3 for discussion.)

Weather variations were controlled for in the analysis by including heating and cooling degree-day variables in the regression model specification.

b. **Background Variables**: Binary variables for various post-audit periods were included in the model to control for the effect of “background” variables. (See Chapter 3, “Regression Model Specification and Variable Definitions” for discussion of the explanatory variables and their interpretation.)

c. **Screening:** See Sections B of this Table 7 and Chapter 2 of the report for description of the data screening for inclusion in the final analysis dataset.

d. **Regression Statistics**: All tables that report the regression parameter estimation results include the coefficient of determination (Adjusted R^2), the t-statistics of all parameter estimates, and the first order auto-regressive parameter value.

e. **Specification:**

1) The model is estimated entirely at the customer level; the sources of variation are cross sectional variations in customer and equipment attributes and temporal variations in weather, as well as time trends.

2) The cooling degree-day and heating degree-day regressors are based on readings of daily high and low temperatures at various locations in the SoCalGas service territory. The bases for the cooling degree-days and heating degree-days are 65 and 70 degrees Fahrenheit, respectively. Other time-dependent regressors are binary variables representing different post audit periods and interactions between degree-days and the binary variables.

3) Self-selection was addressed by matching the distribution of the nonparticipant sample to that of the participant sample with respect to consumption category, weather zone, and service connection period.

4) Various regression model specifications were examined and modified in the process of arriving at the "final" specification. All regression results are reported in Chapter 3.

5) The gross program impacts, combined with effects of "background" trends, were captured in the regression model through binary variables representing the period after the program implementation for both participants and nonparticipants. In some models these variables interacted with the weather variables. The net program impacts were captured in the regression model through binary variables representing participants after the receipt of audits. In some model specifications, these interacted with weather variables.

f. **Error in Measuring Variables:** There was no reason to expect any abnormal measurement error in the data. Many of the variables used are binary and thus not prone to measurement error, which stems from recording errors. Thus, there was no attempt to correct for measurement error in the analysis.

g. **Autocorrelation:** Based on the Durbin-Watson test, first order autocorrelation was found to be present in the dataset used to estimate the regression model parameters. The Yule-Walker method was used to correct for autocorrelation. This method is summarized in the section entitled "Statistical Issues" in Chapter 3. A detailed description of the method and its properties is presented in Greene, 1993.

h. **Heteroskedasticity:** Evidence of some form of heteroskedasticity was found using a White Test (see Greene, 1993, p. 391). The White Correction was used to correct the standard errors of the coefficients computed under OLS. The corrected standard errors are unbiased and consistent. The White test and correction are summarized in the section entitled "Statistical Issues" in Chapter 3.

i. **Collinearity.** Generally, collinearity did not appear to be a problem in the dataset used in the analysis. Various alternative specifications were estimated with fairly consistent results concerning the important coefficient estimates. Most of the t-statistics (corrected) were significant and the estimated coefficients had their expected signs.

This may be taken as evidence that collinearity did not affect the precision of the estimation procedure to any appreciable extent.

j. **Influential Data Points:** Extreme values of the dependent variable (gas consumption per daily) were deleted during the data screening. No influential data diagnostics were performed.

k. **Item Non-Response and Missing Data:** In most cases, any observation with a missing value for one or more variables was eliminated from the regression analysis. The dwelling square feet variable was the one instance where values were imputed for missing responses. This was accomplished by estimating an auxiliary regression. Tests indicated that the estimates of key impact parameters were not sensitive to the inclusion or exclusion of cases with missing values for square feet. (See Chapter 3, "Statistical Issues," for a more detailed discussion.)

l. **Precision:** The standard errors for the savings estimates were calculated from the White covariance matrix of the regression model for the key impact parameters.

m. N/A

n. N/A

5. DATA INTERPRETATION AND APPLICATION

a. **Calculation of Net Impacts** is based on the values of the key regression model impact parameters and the sample means for the participants. This is described in more detail in "Estimates of Net Savings" in Chapter 3.

b**.** **Rationale for Choices:** The process and rationale for choices made in the calculation of net impacts are described in Chapter 3, "Estimation of Net Savings" and "Selection of Best Estimate of Program Savings."