First Year Load Impacts Of Southern California Gas Company's *Program-Year 1996* Commercial Energy Efficiency Incentive Program

Study # 711

Submitted By

Applied Econometrics, Inc.
331 Ninth St.
Del Mar, CA 92014
Voice: (619) 755-1266
Fax: (619) 455-1511
email: parti@inetworld.net

And

Decision Sciences Research Associates
236 West Mountain Street
Pasadena, CA 91103
Voice: (626) 793-9090
Fax: (626) 793-9051
email: chubay@dsra.com

February 28, 1998

Acknowledgments

The people and organizations that contributed data, analyses, and technical expertise to this study are

Southern California Gas Co.
(Project Management, SoCalGas Data Preparation)
James Green
Martin Crundall

AEI
(Econometric Estimation)
Michael Parti
Cynthia Parti

DSRA
(Sampling, Survey Data Collection, Database Management)
Charles Hubay
Kathleen Shields
Alex Currie

MCGI
(SoCalGas Data Preparation)
Kenneth M. Parris

Table Of Contents

Executive Summary

The purpose of this study is to estimate the load impacts of the 1996 Southern California Gas Co. Commercial Equipment Replacement (ER) DSM program on gas cooking in the commercial sector. These cooking impacts are based upon a regression analysis of billing data, and are estimated using the *difference of differences* calculations outlined in Table 5 of the Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Programs[1] (Protocols).

Using the Table 5 approach, we compute the *net* impact of the program by subtracting the change in nonparticipant consumption from the change in participant consumption, after controlling for size and weather differences between participants and nonparticipants. This calculation is performed on a *per measure* basis for cooking. An *ex-post*-net/*ex-ante*-net realization coefficient is also computed. The participant gross impact is computed as the total estimated change in consumption from 1995 to 1997 for the program participants. Realization rates for these impacts are also presented. We compute a *net-to-gross ratio* that is calculated by dividing the net program impact by the simple pre/post program change in participant consumption.

The standard errors for the net effects are relatively straightforward. The standard errors for the gross impacts are somewhat complex, however, because of the nonlinearities in the estimating equation. This means that any gross impact confidence limits would be only approximate. In addition, since the estimated cooking in our model increased from the pre- to the post period, the gross impact is negative, suggesting that the net-to-gross ratio, etc. may be difficult to interpret. Thus we have reported the value of the gross impact but not its confidence interval. Note that even though the estimated *gross* impacts were negative, the estimated *net* impact was positive since the estimated program effects tended to cancel part of the gross impacts for participants. The estimated net impact of the program per installation is slightly greater than the SoCalGas *ex ante* estimates.

The Load Impacts of the SoCalGas Commercial Equipment Replacement Program

The following Table ES.1 summarizes the gross and net load impacts per installation for the cooking end-use category, for the 1996 SoCalGas Commercial Equipment Replacement Program.

[1] As adopted by California Public Utilities Commission Decision 93-05-063; revised January, 1995.

Table ES.1
Average Gross and Net Load Impacts of the 1996 SoCalGas Commercial ER Program

(80% and 90% Confidence Intervals In Parentheses[2])
Units: Annual Therms

Impact Type	*Ex Ante* Estimate Of Cooking Impact *Per Measure*	*Ex Post* Estimate Of Cooking Impact *Per Measure*
Average Gross Impact	998	-201[3]
Average Net Impact	749	793[4] (706 - 975) (674 – 1007)

Realization Rates For Load Impacts Per Measure

Net Realization Rates

The *overall* net realization rate for the SoCalGas ER program is calculated as 1.06. The standard error of this estimate is .13. The 80% confidence interval is (.89 - 1.23).[5] The 90% confidence interval is (.85 - 1.27)[6].

Gross Realization Rates

The *overall* gross realization rate for the SoCalGas ER program is calculated as -.2.[7]

The Net-to-Gross Ratio

The *ex-ante* and *ex-post* net-to-gross ratios for the cooking impacts of the SoCalGas ER Program are shown in the following table.

---

[2] These 80% and 90% confidence intervals are presented in parentheses in each cell, with the 80% intervals above the corresponding 90% interval. See the end of Section IV.A for further discussion of the standard errors.

[3] The average gross impact is negative because the estimated cooking component of the model increased in 1997, the post program period.

[4] This is computed by multiplying the ex-ante net impact estimate by the estimated net impact realization rate.

[5] Note, for the 80% level we use a Z value of 1.28. $Z^*Se = (1.28)^*(.13) \approx .17$

[6] For the 90% level we use a Z value of 1.65. $Z^*Se = (1.65)^*(.13) \approx .21$

[7] This realization rate is negative because the cooking component of the model increased in 1997, the post-program period.

Table ES.2
Net-To-Gross Ratios for the Cooking Load Impacts Of the 1996 SoCalGas Commercial ER Program

Units: Pure Number (A Proportion)

Ex-Ante **Net-to-Gross Cooking**	***Ex-Post*** **Net-to-Gross Cooking**[8]
.75	$-3.95 = \frac{793}{-201} = \frac{ex\ post\ net\ impact}{ex\ post\ gross\ impact}$

[8] This net to gross ratio is negative because the estimated cooking component of the model increased in 1997, the post-program period.

I. Introduction

The focus of the current study is upon estimating the impacts of the 1996 Southern California Gas Co. (SoCalGas) Commercial Equipment Replacement (ER) Program on the use of natural gas for the cooking end-use category in the commercial sector. The basic analytical technique to be used for this impact estimation is a billing regression analysis in which the change in consumption is modeled from the year before the program (1995) to the year after the program (1997). The regression analysis is based upon billing and weather data, survey information, and program tracking information. This approach yields consumption difference estimates which are used to calculate net and gross impact estimates at the end-use level according to the *difference of differences* approach described in Table 5 of the Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Programs[9] (Protocols).

I.A The Program Description

The primary emphasis of the 1996 SoCalGas ER DSM program was upon natural gas savings for cooking. Under this program SoCalGas customers were offered rebates for installing relatively efficient natural-gas-using equipment, primarily, though not exclusively for cooking. The qualifying market segments for this program were SoCalGas commercial customers who were classified as belonging to the following building (business) types. Note, however, that the great majority of participants were in the restaurant category.

1. large office
2. small office
3. restaurant
4. large retail
5. small retail
6. food stores
7. refrigerated warehouses
8. unrefrigerated warehouses
9. hotels and motels
10. primary and secondary education
11. colleges and universities
12. hospitals and clinics
13. miscellaneous commercial

The program was first offered in January 1996 and extended to December 1996.

I.B The Overall Organization Of The Paper

The organization of the paper is as follows. We first briefly describe the overall data set (Section II) and then we discuss the regression results on which the savings estimates are

[9] As adopted by California Public Utilities Commission Decision 93-05-063; revised January, 1995.

based (Section III). After this we present the impact computational methods and the details of the savings impacts for the 1996 SoCalGas ER program (Section IV). The study is summarized in Section V. In Appendix I we discuss the regression specification and estimation procedures. In Appendix II we present the information on EUI's and *ex ante* impacts used in our analysis. Appendix III describes the analysis data base in greater detail.

II. Data Set Summary

The data set used in this analysis is derived from SoCalGas billing records for a set of program participants and nonparticipants. In addition, we make use of recorded and normalized weather data corresponding to the time and place of those billing records. The "normalized" weather data are computed as a 3-year average weather pattern for the weather station corresponding to the billing record. Apart from these basic data, we make use of program participation records for the participants, and survey data.

The dependent variable for the regression equation is the monthly consumption from late 1994 through late 1997. We have up to 45 time-series observations per customer. The original consumption amounts from the billing files have been standardized for the number of days in the billing cycle as follows:

II.1) $Thm_s = (Thm_{bill} / BillDays) \cdot 30.4$

Table II.1
Definitions Of Terms In Equation II.1

Term	Units	Definition
Thm_s	*Thm* per month	*standardized* monthly Therm consumption for a billing month
Thm_{bill}	*Thm* per month	*recorded* Therm consumption for a billing-month
BillDays	days per month	number of days in the billing month
30.4	days per month	30.4 ≈ 365/12 = average number of days per month

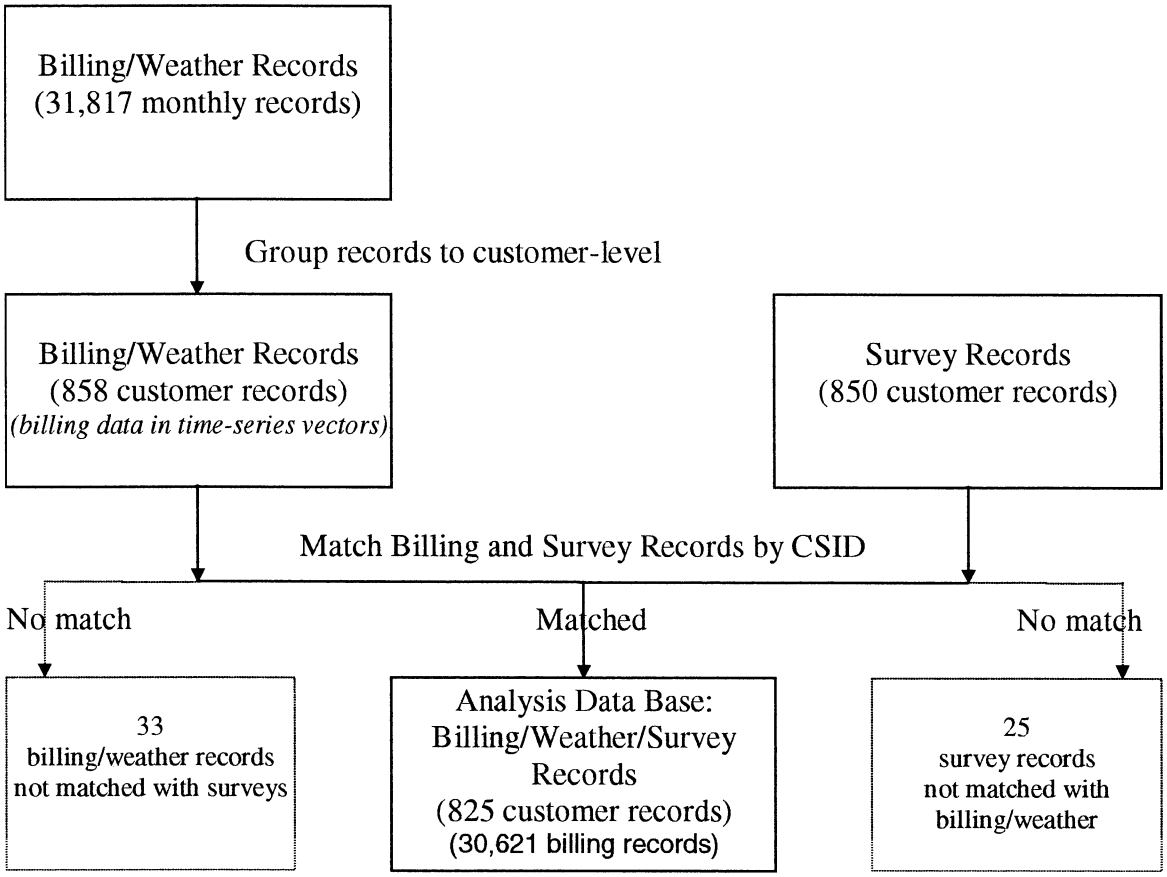
The billing files also contained business-type classification (SIC code) information about the SoCalGas customers analyzed in this study.

The monthly heating-degree-days have been matched to the exact days for each billing month and standardized for the number of days in the billing month using the same approach as described in Equation II.1. In addition, we took account of normal cooling- and heating-degree-days as described in our detailed regression specification discussion in Appendix I.

The analysis data base contains records for 825 SoCalGas customers, including 347 program participants and 478 nonparticipants. Participation status was determined from the SoCalGas program participation files. In the construction of the analysis data base, billing and weather data records were organized into time-series vectors of uniform (45-period) length, and merged with cross-sectional data derived from the telephone survey which was performed for this study. Billing records were available for 858 customers, and survey records were available for 850 customers. The final set of 825 customers includes all customers for whom both survey and non-zero billing data were available.[10] This part of the data flow is shown in Figure II.1 below.

[10] Two of the cases included in the data base failed a later minimum level of gas consumption test and were excluded from the regression analysis. The statistical results are based, therefore, on 823 cases.

Figure II.1
Steps in Constructing the Analysis Data Base
From Billing, Weather and Survey Data Elements



It is important to note that this section of the report details only the last step in a much longer data story. The longer story, which details the derivation of the billing/weather data set from its SoCalGas components and the construction of the original-for-this-study survey data set, is contained in the "Table 7" documentation included at the end of this report. Sampling and matching issues are discussed there, along with an accounting of the surveyed and unsurveyed, participant and nonparticipation populations.

Each customer record in the analysis data base contains the following types of variables:

- a unique case ID,
- SIC code, participation status, weather zone (cross-sectional variables derived from SoCalGas billing and participation files),
- Therm consumption, billing days, and meter-read dates for up to 45 billing intervals (time-series variables derived from SoCalGas billing files)
- Heating degree day data (time-series variables derived from weather files)
- Self-report data on premise and kitchen energy-using equipment (cross-sectional data items derived from the telephone survey performed for this study.)

The full list of variables contained, with definitions and summary statistics, is given in Appendix III.

The preceding diagram detailed the progress of the data set up to the regression analysis. There were further modifications to the data set during the regression analysis as shown in the following table.

Table II
Observations Deleted During Regression Run

Reason For Elimination	Number Eliminated	Number Remaining
Starting Value	-	**30,621**
1. Extremely high value -- appears to be direct meter reading, rather than meter reading difference from last read	1	30,620
2. Therms less than 10 -- appears to be a mistake, beginning or ending partial bill, or pilot light for absent tenant - possibilities not covered by model specification.	869	29,751
3. First nonmissing observation for each remaining customer (deleted for lags used in generalized and first difference estimation forms)	823	28,928
4. Delete observation if lagged value of therms is missing (needed for lags used in difference forms)	39	28,889
Final value in regression		**28,889**

The regression results in the following sections are based upon 823 customers (2 of the original 825 customers are eliminated in step 2) with a total of 28,889 billing period observations.

III. Empirical Results

The object of the regression analysis is to provide a framework for computing a *difference of differences* estimate of the net impact[11] of the 1996 Commercial SoCalGas Equipment Replacement (ER) program. In this section we will first discuss a set of estimating equation issues[12] and then present the results of the regression analysis.

III.A. The Regression Equation Structure

The regression equation is designed to facilitate a comparison of participant vs. nonparticipant consumption changes, taking into account floor space as well as the commercial building type, changes in energy use per square-foot (i.e., changes in overall Energy Use Indices, EUI's[13]), changes in weather, and the presence of installed DSM measures associated with the 1996 SoCalGas ER DSM program. In the current section, we present a summary of the regression equation structure. A detailed discussion of the regression equation is given in Appendix I.

The right side of the equation takes account of factors such as the interaction of weather effects and floorspace, and differing base-period consumption levels for participants and nonparticipants. The regression equation also takes account of the commercial EUI's. These EUI's are the expected energy use per square foot estimates that have been calculated for each interesting end-use/building-type combination by the Forecasting Department of SoCalGas. In this study we have used the SoCalGas estimates of 1992 EUI's.[14]

The basic commercial forecasting model that makes use of EUI's was derived from an early model by Dr. Jerry Jackson[15] that had the following general structure:

$$\text{III.A.1)}\quad \text{Thm}_j = \sum_{bt=1}^{bTypes} \sum_{i=1}^{EUcat} EUI_{i,bt} \cdot FS_{i,bt}$$

where Thm_j is viewed as the sum over all the commercial building types (*bTypes*) and end-use categories (*EUcat*) of the products of the EUI's with the corresponding floor space variables (the FS).

---

[11] The difference of differences approach to net impact calculation is described in some detail in Section IV of this paper.

[12] A detailed discussion of the equation specification is given in Appendix I.

[13] EUI's are discussed in detail in Appendix I.

[14] These are the most recent EUI estimates available.

[15] See, for example, the work by Jerry Jackson, S. Cohn, J. Cope, and W.S. Johnson reported in, "The Commercial Demand for Energy: A Disaggregated Approach," Oak Ridge National Laboratory, ORNL/CON-15, April, 1978. This general orientation is still being used by the CEC and many utilities and is the basis of the EPRI COMMEND commercial end-use forecasting model.

In our regression equation, we take account of the building type in order to use the EUI's to obtain a prior estimate of the expected usage. To obtain the expected consumption for a customer, we multiply each EUI by the floor space. These expected consumption levels are modified by the estimated regression coefficients to account for differences between the expected and actual consumption levels. In addition, the regression is designed to account for the consumption patterns associated with customers who did not furnish information about their floorspace, as well as adjustments for business activity levels and changes in the stock of cooking equipment.

The structure of our regression equation is as follows. The dependent variable is the standardized[16] monthly consumption from 1995 through 1997. The explanatory variables for this regression include program participation dummy variables, heating degree-days, prior program impact estimates, indicators of actual and missing floor space, variables indicating the program year (1996) and the impact measurement year (1997), and variables accounting for differences in realization rates of prior EUIs across building types and across end-use categories. Our estimation effort was directed toward a measurement of program impacts that persisted into 1997. Our *particular* goal was to test whether there is any evidence of overall SoCalGas Commercial DSM program impacts after controlling for changes in consumption that may have been due to more general factors in the impact measurement year.

In our regression equation we estimate coefficients for the EUI terms that account for differing overall consumption practices by participants and nonparticipants as well as general changes in overall consumption, accounting for program-related impacts. Thus, our model estimates adjusted EUI coefficients for the base year, program year and impact year. We tested whether these EUI coefficients differed across customer categories and also whether the end-use components of the overall EUIs should be given different weights. These EUIs were designed for average customers and we also tested whether they should be modified for the upper and lower tails of the floorspace distribution.[17] Finally, we computed a realization rate for the SoCalGas prior estimates of net program savings. We can write the form of the basic regression equation specification as follows:

$$\begin{aligned} Thm = {} & (AdjNpYr)\cdot\sum_{bt=1}^{BT}\alpha_{bt}\cdot d_{bt}\cdot EUIFS_{bt} \\ & +\gamma_{sh}\cdot EUI_{sh,rest}+\gamma_{nonSh}\cdot EUI_{nonSh,rest} \\ & +\phi_{sh}\cdot dfEUI_{sh,bt}+\phi_{nonSh}\cdot dfEUI_{nonSh,bt} \\ & +(r+\theta\cdot AddEquip)\cdot(M\cdot d_{part}\cdot d_{97})+e. \end{aligned} \qquad \text{III.A.2)}$$

where,

---

[16] The billing-day standardization procedure is described above in Section II.

[17] The *average* EUI has been studied traditionally because of its importance for forecasting average annual loads. The EUIs may differ, however, across the floorspace distribution and these differences are important information for DSM program design as well as possible target marketing by competitive energy companies. We have investigated the upper and lower 1% tails of the floorspace distribution in the present study. The results are encouraging and indicate that a wider investigation, beyond the scope of the present work, might be an important study in itself.

III.A.3) $$AdjNpYr = 1 + \beta_{npart} \cdot d_{npart} + (\beta_{96} + \beta_{part,96} \cdot d_{part}) \cdot d_{96} + \beta_{97} \cdot d_{97}$$

III.A.4) $$EUIFS_{bt} = \left[1 \cdot \left(\sum_{i=sh}^{o} EUI_{i,bt} \cdot S_i \right) + a \cdot EUI_{sh,bt} S_{sh} + c \cdot EUI_{ck,bt} S_{ck} \right] \cdot FSterm$$

III.A.4a) $$a = a_{sh} + a_{sh,97} \cdot d_{97}$$

III.A.5) $$\begin{aligned} c = {} & a_{ck} + (a_{ck,HighVolume} \cdot dHighVolumeMonth) \\ & + [a_{ck,97} + (a_{ck,97,IncreasedActivity} \cdot dActivityUp) \\ & + (a_{ck,97,decreasedActivity} \cdot dActivityDown)] \cdot d97 \end{aligned}$$

III.A.6) $$FSterm = FSspec + d_{miss} + d_{10k}$$

III.A.7) $$dfEUI_{i,bt} = \frac{EUI_{i,bt} \cdot S_{i,bt} - EUI_{i,rest} \cdot S_{i,rest}}{EUI_{i,rest} \cdot S_{i,rest}} ; i = sh, ck, o$$ [18]

In these equations, the customer category subscript index *bt* refers to the building (business) type[19], and *BT* (in upper case) represents the number of building types; *npart* means nonparticipant and *part* means participant; the end-use indices i = *sh,ck,o* refer to the space-heating, cooking and 'other' end-use categories; and the subscript *nonSH* refers to the cooking and 'other' end-use categories. The α_{bt} are the overall adjusted EUI coefficients for the base year (1995); and a_{sh} and a_{ck} represent incremental EUI coefficients for the space heating and cooking end-use categories.

Dummy variables d_{96} and d_{97} have the value 1 if the current consumption year is the one that is indicated (they are 0 otherwise); and the corresponding coefficients, the β's, are the proportionate overall general differences in consumption from the base year (1995) to the indicated year (1996 or 1997).

The overall *multiplicative* adjustment to the base year (1995) EUI's is given by the terms in which the α's and the β's appear. In addition, we account for a general *additive* adjustment to the overall EUI's, apart from floor space and the other interaction terms, with the γ and ϕ parameters.

The symbol *M* is used to represent the *ex ante* estimate of the net impact per installation of measures associated with the cooking end-use in the ER program. These prior program net impact estimates are zero for nonparticipants. The *r* term is a net savings realization coefficient.

---

[18] Note that $dfEUI_{nonSh,bt} = dfEUI_{ck,bt} + dfEUI_{o,bt}$

[19] See Appendix A.II.1 for a listing of the building types.

The FS variable in Equation III.A.2 represents floor space, and the dummy variables, d_c, are indicators of program participation ($c = p$) or nonparticipation ($c = n$). The term d_i is a dummy variable indicator of participation with respect to measure i (i = space heating, cooking and 'other') in the ER program. Finally, e is an additive error term for the equation.

In addition to this brief introduction, a detailed discussion of this estimating equation is offered to the interested reader in Appendix I, along with a description of our estimation procedure. Apart from this, we have included a brief explanation of each model parameter in the following table of regression results.

III.B The Results Of The Regression Analysis

The following table contains the results of the regression analysis. The first column contains the terms as they appear in the discussion of the overall equation specification in the previous subsection (Equation III.A.2) and in Appendix I.

Table III.1
Regression Results

Dependent Variable: Standardized Therms per customer per month (*Thm*)
$R^2 = .67$
Number Of Observations =28,889

Symbol In Equations III.A.2 - III.A.7	Symbol Definition	Regression Parameter Estimates (t-ratios)
r	realization coefficient for *ex ante* estimates of net impacts associated with the 1996 SoCalGas ER DSM program	-1.06 (-8.23)
θ	increment to net impact realization coefficient to account for additional, rather than replacement, equipment installed in connection with program.	0.33 (1.44)
β_{npart}	overall difference (roughly proportionate) in consumption between nonparticipants and participants.	-0.48 (-106.09)
β_{96}	overall change (roughly proportionate) in consumption from the pre-program period to the program year, 1996.	-0.04 (-6.90)
$\beta_{part,96}$	increment to overall change in consumption from the pre-program period to 1996 for participants. This may result from program-related effects during the program year. These effects are not, however, included in our calculation of pre-post impact differences.	-0.02 (-2.41)
β_{97}	overall change (roughly proportionate) in consumption from the pre-program period to the post-program year 1997.	-0.12 (-17.07)
$\alpha_{restaurant}$	overall adjusted EUI coefficient for the base year (1995): restaurant category	0.63 (21.81)
$\alpha_{foodStore}$	overall adjusted EUI coefficient for the base year (1995): food store category	3.19 (22.29)
$\alpha_{K\text{-}12\ schools}$	overall adjusted EUI coefficient for the base year (1995): K-12 school category	25.19 (51.13)

Dependent Variable: Standardized Therms per customer per month (*Thm*) $R^2 = .67$ *Number Of Observations =28,889*		
Symbol In Equations III.A.2 - III.A.7	**Symbol Definition**	**Regression Parameter Estimates (t-ratios)**
$\alpha_{healthFacilities}$	overall adjusted EUI coefficient for the base year (1995): health-facilities category	3.42 (40.54)
$\alpha_{miscellaneous}$	overall adjusted EUI coefficient for the base year (1995): miscellaneous building-type category	3.98 (23.56)
$\alpha_{FS\ upper\ 1\%}$	increment to EUIs for customers in upper 1% of floorspace distribution	-0.47 (-13.54)
$\alpha_{FS\ lower\ 1\%}$	increment to EUIs for customers in lower 1% of floorspace distribution	3.93 (17.05)
$\alpha_{restaurant,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): restaurant category, observations with missing floorspace	2313.74 (21.72)
$\alpha_{K\text{-}12\ schools,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): K-12 school category, observations with missing floorspace	78955.32 (54.23)
$\alpha_{healthFacilities,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): health-facilities category, observations with missing floorspace	35525.71 (130.21)
$\alpha_{hotel/motel,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): hotel/motel category, observations with missing floorspace	132629.68 (39.34)
$\alpha_{miscellaneous,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): miscellaneous building-type category, observations with missing floorspace	49640.64 (43.78)
$\alpha_{colleges,\ floorspace>10,000\ sqft}$	overall adjusted EUI coefficient for the base year (1995): college category, observations with floorspace greater than 10,000 square feet	173210.74 (5.70)
$\alpha_{healthFacilities,\ floorspace>10,000\ sqft}$	overall adjusted EUI coefficient for the base year (1995): health-facilities category, observations with floorspace greater than 10,000 square feet	385304.17 (45.25)
$\alpha_{hotel/motel,\ floorspace>10,000\ sqft}$	overall adjusted EUI coefficient for the base year (1995): hotel/motel category, observations with floorspace greater than 10,000 square feet	162051.07 (11.23)
a_{sh}	incremental EUI coefficients for the space heating end-use category.	-0.37 (-25.85)
$a_{sh,97}$	incremental EUI coefficients for the space heating end-use category in the post period, 1997.	0.09 (4.87)
a_{ck}	incremental EUI coefficients for the cooking end-use category.	1.11 (7.49)
$a_{ck,97}$	incremental EUI coefficients for the cooking end-use category in 1997.	0.59 (7.28)
$a_{ck,97,\ increased\ activity}$	incremental EUI coefficients for the cooking end-use category in 1997 for customers who claimed increased volume of activity in their establishments.	0.33 (5.07)
$a_{ck,97,\ decreased\ activity}$	incremental EUI coefficients for the cooking end-use category in 1997 for customers who claimed decreased volume of activity in their establishments.	-0.32 (-4.23)
$a_{ck,HighVolume}$	incremental EUI coefficients for the cooking end-use category in the high volume months identified by respondents.	0.38 (6.75)

Dependent Variable: Standardized Therms per customer per month (*Thm*) R^2 = .67 *Number Of Observations =28,889*		
Symbol In Equations III.A.2 - III.A.7	**Symbol Definition**	**Regression Parameter Estimates (t-ratios)**
γ_{sh}	general *additive* adjustment to the space heating EUI, apart from floor space and the other interaction terms (the restaurant EUI is used for this general purpose)	3014.27 (23.23)
γ_{nonSh}	general *additive* adjustment to the nonspace heating EUI, apart from floor space and the other interaction terms (the restaurant EUI's are used for this general purpose)	3331.60 (161.49)
ϕ_{sh}	incremental *additive* adjustment to the space heating EUI accounting for differences between the customer's business-type EUI for space heating and the restaurant EUI for space heating, apart from floor space and the other interaction terms	-14.11 (-12.17)
ϕ_{nonSh}	incremental *additive* adjustment to the nonspace heating EUI accounting for differences between the customer's business-type EUI for nonspace heating end-uses and the restaurant EUIs for the nonspace heating end-uses, apart from floor space and the other interaction terms	20.00 (6.87)

As indicated by this table, there is general agreement between the results and the prior estimates of the program impacts. The net impact realization coefficient (r) is *-1.06* with a relatively high (in absolute value) t-value (t_r = - 8.2). The estimate of the θ parameter is positive, as expected, since it accounts for additions to (rather than replacements of) the cooking appliance stock in connection with the SoCalGas program. The t-ratio of this variable is, however, small.

The base year adjusted consumption parameter, β_{npart}, for nonparticipants is -.48. This suggests that, other things equal, consumption for participants tends to be greater than consumption for nonparticipants. Note that since we included the γ and ϕ terms in the regression, this coefficient should not be interpreted as a -48% increment to an overall EUI realization rate.

The negative signs of the β_{96}, $\beta_{part,96}$ and β_{97} parameters indicate that, other things equal, consumption in 1996 and 1997[20] generally fell for all customers in comparison with the pre-program levels of consumption. The sign of the $\beta_{part,96}$ parameter indicates that the consumption level of participants fell by more than the consumption level of nonparticipants during the program year. We have not ascribed this difference to the program in our impact calculations, since we wished to have a clear definition of the pre- and post-program periods that would apply identically to the participants and nonparticipants.

The pre/post definition used in our program impact calculations (see section IV) refers to pre- and post-program year, rather than the precise participant installation date. Use of

[20] 1996 and 1997 are the program year and post-program year respectively.

the precise installation date for participants would beg the difficult question of defining the comparable date to use for nonparticipants. Of course, to the extent that there is any decline in the impact of the cooking measures over time, our procedure would tend to bias the calculated impacts downward, resulting in estimates of program effects that are too conservative.

The negative signs of the β_{96} and β_{97} parameters apply to nonparticipants as well as participants and might be due to general trend effects or to a general program spillover effect. We have not structured the current analysis to distinguish between these.

The α parameters are included in the analysis to account for variations in EUIs across building(business)-types and other customer categories. Note, that these parameters should not be regarded as EUI base realization rates because of the existence of the γ and ϕ terms in the regression equation. The α parameters do, however, account for customer-to-customer differences in the energy use per square foot variations across customers that are not accounted for by the prior EUIs that were furnished by the SoCalGas forecasting department. The differences in these coefficients across the indicated categories (restaurants, ..., miscellaneous) are affected by cross-category differences between the actual energy use per square foot and the prior EUIs. In addition, it is possible that they would be affected by differences in the accuracy of the claimed customer floorspace estimates across the business-type categories. Note that it is possible for these coefficients to be positive or negative because the customer consumption levels are partly accounted for by the γ and ϕ terms.

The incremental α parameters for variation in the EUI over the floorspace distribution are negative for the very large floorspace and positive for the very small floorspace.

The coefficients $\alpha_{restaurant}$ through $\alpha_{miscellaneous}$ refer to observations for which we have information on the floorspace supplied by survey respondents. The coefficients

$$\alpha_{restaurant,\ missing\ floorspace} \text{ through } \alpha_{miscellaneous,\ missing\ floorspace}$$

refer to observations for which we have missing floorspace line items on our questionnaire responses. For respondents in this latter category, the estimated coefficients account for differences in floorspace as well as differences in the applicability of the prior EUIs.

The coefficients $\alpha_{colleges,\ floorspace>10,000\ sqft}$ through $\alpha_{hotel/motel,\ floorspace>10,000\ sqft}$ refer to observation for which the respondents furnished us with information that the floor space is greater than 10,000 square feet. In these cases, however, the precise floorspace was not provided. Thus we treated these observations analogously to the observations with completely missing information about the floorspace, and the estimated coefficients account for differences in floorspace as well as differences in the applicability of the prior EUIs.

In general, these α parameters are positive. In this connection, it should be noted that these parameters refer to only a part of predicted consumption and that none of the

predicted consumption levels are negative.

The a parameters were used to test whether the end-use components of the overall EUIs should be given different weights. To investigate these questions, we divided the equipment into three broad end-use category: space heating, cooking and other. We chose this categorization because space heating is almost always an important source of variation in natural gas consumption, and cooking was the focus of the 1996 SoCalGas Equipment Replacement program.

The a_{sh} parameter is negative, indicating that the space heating component of the $EUIFS_{bt}$ term should have a lower weight than the 'other' EUI component.[21] Note, however, that this coefficient is slightly less negative in 1997 (the coefficient $a_{sh,97}$ is negative). Because of the particular importance of the cooking end-use category in our study, we tested for the effects of several variables on cooking.[22] The adjustment to the cooking EUI can be viewed as the linear function shown in equation III.A.5. All the estimated parameters of this function are generally positive, indicating that the weight for cooking in the $EUIFS_{bt}$ equation generally would be higher than the 'other' EUI component. There is, however a negative sign for the $a_{ck,97,decreasedActivity}$ parameter that accounts for business that reported a decline in the volume of activity.

The a_{ck} parameter (>0) might be viewed as the intercept of equation III.A.5.[23] The $a_{ck,HighVolume}$ parameter is included to test whether the respondents' claimed high volume months were reflected in higher EUIs for cooking.[24] The positive sign of this parameter indicates general agreement with the claims of the survey respondents. Because of the importance of cooking to the SoCalGas 1996 ER program, we wished to account for any general trend influences in the post-program period (1997) that were related particularly to cooking. For this reason, we included the $a_{ck,97}$ term in the equation. The $a_{ck,97}$ parameter is positive, indicating that the cooking portion of the load should have a higher weight during the post-program period.

The remaining γ and ϕ parameters were included to account for any general *additive* adjustment to the overall EUI's, apart from floor space and the other interaction terms. For this purpose, we divided the end-use equipment into space heating and all-other categories. This part of the analysis was carried out in two logical steps: 1. a step to see

---

[21] The 'other' component has a weight of '1' in the $EUIFS_{bt}$ terms shown in equation III.A.4 because the EUI for the other category appears only in the expression $\left[1 \cdot \sum_{i=sh}^{o} EUI_{i,bt} \cdot S_i\right]$ in the $EUIFS_{bt}$ term. By contrast, the weight for the space heating EUI in this term would be $1 + a_{sh}$, and the corresponding weight for the cooking EUI would be $1 + c$.

[22] See equations III.A.4 and III.A.5.

[23] Of course, equation III.A.5 is not a standalone regression equation. It is only a detailed expansion of the c term in equation III.A.4, which is itself a component of equation III.A.2. As noted above, the complete equation is a nonlinear function which is estimated using iterative numerical techniques.

[24] The surveyed customers were asked, "During which months do you experience the heaviest volumes?" The variable *dHighVolumeMonth* has the value 1 for each month that was claimed as a high volume month and is zero otherwise.

if a standard set of EUIs for restaurants would suffice to account for an adjustment to the EUIs apart from the other EUI interaction terms;[25] and, 2. a step to test whether proportionate differences between each observation's EUIs and the restaurant EUIs were useful adjustments to the EUIs after taking into account the other interactions (floorspace, etc.). All but one of these terms were positive, with large (in absolute value) t-ratios. Only the proportionate difference term for space heating EUIs was negative.

In general, the t-ratios for the important variables are high and the net savings realization coefficient is reasonably consistent with our prior information. The estimation method used[26] accounted for the serial correlation and heteroskedasticity issues. To obtain the serial correlation coefficient, we made use of an overall sample-wide expected value estimated by regressing the current residual against the lagged residual. To obtain the weights used in connection with the heteroskedasticity issue, we estimated an error variance equation (after applying the serial correlation coefficient). In the latter equation, the cross-sectional portion of the error variance was modeled as a semi-log function of the floor space, the program participation status of the customer, and variables indicating whether the customer had a missing floorspace response or the floorspace was in excess of 10,000 square feet. The details of these equation error models are presented in Appendix I.

[25] Note that the effect of adding the nonspace-heating portion of the restaurant EUI term is to add a constant term to the equation.

[26] Nonlinear least squares. See Appendix I for a description of this procedure.

IV. Savings Estimates

The savings estimates for the 1996 SoCalGas ER program were generated using the *difference of differences* approach outlined in Table 5 of the Protocols. In this section we first review, briefly, the impact calculations recommended in Table 5 of the Protocols and then present a set of tables containing our calculations.

IV.A The Impact Calculation Methods

In the approach described in Table 5 of the Protocols, the change in comparison group consumption from the pre-program to the post-program period is viewed as representing the corresponding change in consumption for participants that would have occurred in the absence of the program. According to this view, therefore, we would compute the *net* impact of the program by subtracting the change in comparison group consumption from the change in participant consumption. This can be summarized as follows:

IV.1) $$Net\ Impact = (Thm_{pre,\,part} - Thm_{post,\,part}) - (Thm_{pre,\,npart} - Thm_{post,\,npart})$$

Table IV.A.1
Definitions Of Terms In Equation IV.1

Term	Units	Definition
$Thm_{pre,g}$; *g = Program-Group, Comparison-Group*	Annual *Thm*	annual Therm consumption during the *pre-program* period (the *base* period) by group *g* (participant or comparison group)
$Thm_{post,g}$; *g = Program-Group, Comparison-Group*	Annual *Thm*	annual Therm consumption during the *post-program* period by group *g* (participant or comparison group)
part	index	Program Group
npart	index	Nonparticipant Comparison Group

Using a more compact notation, we can rewrite this equation as

IV.2) $$Net\ Impact = \Delta Thm_{part} - \Delta Thm_{npart}$$

Table IV.A.2
Definitions Of Terms In Equation IV.2

Term	Units	Definition
ΔThm_g; *g = Program-Group(part), Comparison Group(npart)*	Annual *Thm*	$Thm_{pre,g} - Thm_{post,g}$

In this context, *for similar participants and nonparticipants*, the change in the nonparticipant consumption is assumed to equal the change in participant consumption in

the absence of the program. To clarify this point, we can rewrite equation IV.2 in conceptual terms as follows:

IV.2a) $Net\ Impact = Gross\ Impact_{part} - Common\ Impact_{Refers\ to\ Both\text{-}Groups}$

The gross impact for the program group consists of the simple change in consumption for the participants from the base period (1995 in our case) to the impact evaluation period (1997 in our case). The common impact, as noted above, can arise from trend-related economic effects as well as free-rider and other effects. Since our regression equation takes into account the fact that customers differ in floorspace, weather and other characteristics, we can control for these factors in calculating the program impacts. Note that, after controlling for these factors, if this common impact term did not apply equally to participants and the comparison group, it would not be appropriate to subtract that term from the participant gross impact to compute the net impact.

To show how these issues bear upon our study, we first reproduce our regression equation specification from section III for the convenience of the reader.

III. A.2)
$$\begin{aligned} Thm = {} & (AdjNpYr)\cdot\sum_{bt=1}^{BT}\alpha_{bt}\cdot d_{bt}\cdot EUIFS_{bt} \\ & +\gamma_{sh}\cdot EUI_{sh,rest}+\gamma_{nonSh}\cdot EUI_{nonSh,rest} \\ & +\phi_{sh}\cdot dfEUI_{sh,bt}+\phi_{nonSh}\cdot dfEUI_{nonSh,bt} \\ & +(r+\theta\cdot AddEquip)\cdot(M\cdot d_{part}\cdot d_{97})+e. \end{aligned}$$

where,

III.A.3) $$AdjNpYr = 1+\beta_{npart}\cdot d_{npart}+(\beta_{96}+\beta_{part,96}\cdot d_{part})\cdot d_{96}+\beta_{97}\cdot d_{97}$$

III.A.4) $$EUIFS_{bt} = \left[1\cdot\left(\sum_{i=sh}^{o} EUI_{i,bt}\cdot S_i\right)+a\cdot EUI_{sh,bt}S_{sh}+c\cdot EUI_{ck,bt}S_{ck}\right]\cdot FSterm$$

III.A.4a) $$a = a_{sh}+a_{sh,97}\cdot d_{97}$$

III.A.5)
$$\begin{aligned} c = {} & a_{ck}+(a_{ck,HighVolume}\cdot dHighVolumeMonth) \\ & +[a_{ck,97}+(a_{ck,97,IncreasedActivity}\cdot dActivityUp) \\ & +(a_{ck,97,decreasedActivity}\cdot dActivityDown)]\cdot d97 \end{aligned}$$

III.A.6) $$FSterm = FSspec + d_{miss} + d_{10k}$$

III.A.7) $$dfEUI_{i,bt} = \frac{EUI_{i,bt}\cdot S_{i,bt}-EUI_{i,rest}\cdot S_{i,rest}}{EUI_{i,rest}\cdot S_{i,rest}};\ i = sh,\ ck,\ o$$

Using these equations (but ignoring the error term), we express the base year consumption for participants as follows:

$$Thm_{part,95} = (AdjNpYr_{part,95}) \cdot \sum_{bt=1}^{BT} \alpha_{bt} \cdot d_{bt} \cdot EUIFS_{bt,95}$$
$$+ \gamma_{sh} \cdot EUI_{sh,rest} + \gamma_{nonSh} \cdot EUI_{nonSh,rest} \qquad \text{IV.3a)}$$
$$+ \phi_{sh} \cdot dfEUI_{sh,bt} + \phi_{nonSh} \cdot dfEUI_{nonSh,bt}$$

or,

$$Thm_{part,95} = (1) \cdot \sum_{bt=1}^{BT} \alpha_{bt} \cdot d_{bt} \cdot EUIFS_{bt,95}$$
$$+ \gamma_{sh} \cdot EUI_{sh,rest} + \gamma_{nonSh} \cdot EUI_{nonSh,rest} \qquad \text{IV.3b)}$$
$$+ \phi_{sh} \cdot dfEUI_{sh,bt} + \phi_{nonSh} \cdot dfEUI_{nonSh,bt}$$

where $AdjNpYr_{part,95} = 1$ is the value of $AdjNpYr$ for participant customers in 1995, and $EUIFS_{bt,95}$ is the value of $EUIFS_{bt}$ for 1995.[27]

The base year consumption for nonparticipants would be expressed as follows:

$$Thm_{npart,95} = (1 + \beta_{npart}) \cdot \sum_{bt=1}^{BT} \alpha_{bt} \cdot d_{bt} \cdot EUIFS_{bt,95}$$
$$+ \gamma_{sh} \cdot EUI_{sh,rest} + \gamma_{nonSh} \cdot EUI_{nonSh,rest} \qquad \text{IV.4)}$$
$$+ \phi_{sh} \cdot dfEUI_{sh,bt} + \phi_{nonSh} \cdot dfEUI_{nonSh,bt}$$

Note that we take account of the participant/nonparticipant differences in the base term calibration coefficient (β_{npart}). Having accounted for these differences, however, we estimate a common change in consumption from 1995 to 1997 that applies to both groups.

Thus, accounting only for the 1997 impact adjustment terms that are common to both participants and nonparticipants, and ignoring the 1997 *program-related* term for participants, we can write the 1997 participant consumption equation as:

[27] Note that the term $EUIFS_{bt}$ is not a function of customer participation status; thus, it is the same for participants and nonparticipants. $EUIFS_{bt}$ depends upon the year because of the ack,97 term (see equation III.A.5).

$$Thm_{part,97} = (1+\beta_{97}) \cdot \sum_{bt=1}^{BT} \alpha_{bt} \cdot d_{bt} \cdot EUIFS_{bt,97}$$

IV.5)
$$+\gamma_{sh} \cdot EUI_{sh,rest} + \gamma_{nonSh} \cdot EUI_{nonSh,rest}$$
$$+\phi_{sh} \cdot dfEUI_{sh,bt} + \phi_{nonSh} \cdot dfEUI_{nonSh,bt}$$

Again, accounting only for the factors common to participants and nonparticipants, then, we can write the *change* in consumption from 1995 to 1997, for participants, as follows:

IV.6a) Common Impact = $Thm_{part,95}$ - $Thm_{part,97}$

or,

$$Common\ Impact = \sum_{bt=1}^{BT} \alpha_{bt} \cdot d_{bt} \cdot EUIFS_{bt,95}$$

IV.6b)
$$-(1+\beta_{97}) \cdot \sum_{bt=1}^{BT} \alpha_{bt} \cdot d_{bt} \cdot EUIFS_{bt,97}$$

Simplifying this, we have,

IV.6c) $$Common\ Impact = \sum_{bt=1}^{BT} \left[\alpha_{bt} \cdot d_{bt} \cdot \left(EUIFS_{bt,95} - (1+\beta_{97}) \cdot EUIFS_{bt,97}\right)\right]$$

Limiting our attention to the cooking portion of the load we can write

IV.7) $$Common\ Impact_{ck} = \sum_{bt=1}^{BT} \left[\alpha_{bt} \cdot d_{bt} \cdot \left(EUIFS_{ck,bt,95} - (1+\beta_{97}) \cdot EUIFS_{ck,bt,97}\right)\right]$$

We have accounted for the common impact term in our regression equation. We also account (with the θ term in equation III.2) for the increase in the stock of cooking equipment bought by participants. In addition, we include a term for participants that accounts for the incremental, *program-related*, change in participant consumption. We write this *net impact* term as follows:

IV.8) $$NetImpact = -r \cdot (M \cdot d_{part} \cdot d_{97}).$$

Rearranging equation IV.2a we compute the gross impact as follows:

IV.9) $Gross\ Impact_{part}$= $Net\ Impact_{part}$ + *Common Impact*

We also present this net impact information in a somewhat different form by calculating a ratio of the net program impact to the gross program impact, the *net to gross* ratio, as follows:

IV.10) $$Net\,To\,Gross\,Ratio = \frac{Net\ Impact}{Gross\ Impact}$$

This *Net-To-Gross Ratio* was originally designed to facilitate the discussion of the relative sizes of the net program impact and other changes in consumption such as the free-rider effect.

The standard errors for the net effects are relatively straightforward. The standard errors for the gross impacts are somewhat complex, however, because of the nonlinearities in the estimating equation. Note, for example, the implied product $\alpha_{bt} \cdot \beta_{97}$ in equation IV.6c. This means that any gross impact confidence limits would be only approximate. In addition, since cooking increased from the pre- to the post period, the gross impact is positive, making the usual interpretation of free-ridership, etc. difficult to interpret. Thus we have reported the value of the gross impact but not its confidence interval.

To summarize these calculations, the net impact confidence intervals were calculated using the standard approach. The gross impact level, but not confidence intervals were reported. Note, however, that all the t-ratios involved are relatively large.

IV.B The Impact Calculations

The tables in this section contain the results of the impact calculations for this study. These are:

1. the gross and net load impacts (annual Thm/installation) for the ER Program;
2. the overall net realization rate, and the gross realization rate for the ER program; and
3. the net-to-gross ratios for the ER program.

The Load Impacts of the ER Program

Table IV.B.1 summarizes the load impacts of the 1996 SoCalGas Commercial Equipment Replacement (ER) Program. These impacts are calculated for two categories of end-use, on a measure basis for cooking, and in terms of annual Therms, and using normalized weather, for HVAC.

Table IV.B.1
Average Gross and Net Load Impacts of the
1996 SoCalGas Commercial ER Program
(80% and 90% Confidence Intervals In Parentheses[28])
Units: Annual Therms

Impact Type	*Ex Ante* Estimate Of Cooking Impact *Per Measure*	*Ex Post* Estimate Of Cooking Impact *Per Measure*
1. Avg. Gross Impact	998	-201[29]
2. Avg. Net Impact	749	793[30] (706 - 975) (674 – 1007)

Net and Gross Realization Rates For the ER Program

Net Realization Rates

The *overall* net realization rate for the SoCalGas ER program is calculated as 1.06. The standard error of this estimate is .13. The 80% confidence interval is (.89 - 1.23).[31] The 90% confidence interval is (.85 - 1.27)[32].

---

[28] These 80% and 90% confidence intervals are presented in parentheses in each cell, with the 80% intervals above the corresponding 90% interval. See the end of Section IV.A for further discussion of the standard errors.

[29] Note, the gross impact is negative because the cooking component of the model increased in 1997, the post program period.

[30] This is computed by multiplying the ex-ante net impact (749) estimate by the estimated net impact realization rate (1.06).

[31] Note, for the 80% level we use a Z value of 1.28. $Z^*Se = (1.28)^*(.13) \approx .17$

[32] For the 90% level we use a Z value of 1.65. $Z^*Se = (1.65)^*(.13) \approx .21$

Gross Realization Rates

The following Table IV.B.2 shows the realization rates calculated for the gross load impacts per installation, for the ER Program.

Table IV.B.2
Average Gross Realization Rates for the Load Impacts Per Installation
For the 1996 SoCalGas ER Program

Units: Pure Number (Ratio)

Realization Rate Cooking[33]
$-.2 = \frac{-201}{998} = \frac{ex\ post\ gross\ impact}{ex\ ante\ gross\ impact}$

Table IV.B.3 below shows the *ex post* net-to-gross ratio.

Table IV.B.3
***Ex Post* Net-To-Gross Ratio for the Load Impacts**
Of the 1996 SoCalGas Commercial Equipment Replacement Program
Calculated for Load Impacts Per Installation For Cooking
Units: Pure Number (A Proportion)

Net-to-Gross, Cooking[34]
$-3.95 = \frac{793}{-201} = \frac{ex\ post\ net\ impact}{ex\ post\ gross\ impact}$

[33] Again, this realization rate is negative because the cooking component of the model increased in 1997, the post-program period. This realization rate is computed as the *ex post* gross impact/*ex ante* gross impact = -201/998.

[34] This net to gross ratio is negative because the cooking component of the model increased in 1997, the post-program period.

V. Summary

In this study we have investigated the impacts of the 1996 SoCalGas Commercial Equipment Replacement program. The study is based upon a regression analysis of pre/post-program changes in consumption taken from monthly billing data. The regression analysis takes account of the size of the firms (floor space), normalized and actual weather, the building-type classification, the participation status for the 1996 SoCalGas ER program and the energy use indices taken from the SoCalGas forecasting group.

Using the regression equation and information about a sample of commercial customers, we are able to estimate the net program impacts using the Difference-Of-Differences approach outlined in Table 5 of the Protocols.

The estimated program impacts for cooking are presented on a measure level. In addition to the net program impacts, we also present a net-to-gross impact ratio for each program element, as well as a set of gross and net impact realization rates.

Apart from this equation we also estimated our regression equations using an ordinary least squares approach, and a first difference approach.[35] All of these results are summarized in Appendix I. For each of these estimation approaches, the program impact estimate was larger than the estimates presented in the current section. The ordinary least squares (OLS) approach gives unbiased coefficients but it is inefficient, since further information about the error structure may aid the estimation process.

The first difference approach has substantially the same form as the equation presented in the section III (weighted, generalized difference) and the OLS form. The first difference form is somewhat different, however, in that any term that is constant for each individual customer would cancel. There are only two terms for which this would be true for the current equation: the nonspace-heating terms associated with the parameters γ_{nonSh} and ϕ_{nonSh}. Of course, there could be other terms that are not in the current model that would also be cancelled by such a procedure. Thus, there are other models that would be consistent with our first-difference regression equation. In this sense, we might think of the first-difference regression as less constrained than the generalized difference form used to generate the parameters shown in table III.1.

A related issue has to do with the question of self-selection bias and whether there are participant characteristics that are correlated with the error term. A selectivity correction term could be used to address this issue and would be modeled as a function of cross-customer, not time-series, characteristics. This means that such a term would cancel and thus be unnecessary in our first-difference regression. Since the first-difference regression was therefore corrected for any self-selection bias and since it resulted in a larger estimate of the program impact coefficient, there seemed to be no compelling

[35] Note, the same number of observations was used in each of these estimation approaches since we eliminated the first observation for each customer no matter which approach was used.

argument to include a selectivity correction term in the regression model presented in section III.

A general question of the applicability of the models arises. How do we choose between the models? The model presented in section III is corrected for serial correlation and heteroskedasticity. Although there are some advantages over the OLS model, the OLS model does not rely on large samples for its attractive properties. By contrast, the current model can claim consistency although we are a long way from an infinite sample size since we are dealing with 823 firms with an average of approximately 35 time-series observations for each firm. Nonetheless, an assortment of modeling techniques does give us an idea of the sensitivity of the savings estimate to the type of technique used. In the current impact calculations we have presented the results obtained from the model in which we address the serial correlation and heteroskedasticity issues. This model gave the most conservative estimate for the net impact realization rate.

Appendix I: Regression Specification And Estimation Issues

In this appendix, we first discuss the specification of our regression equation. After this, we describe the nonlinear estimation procedure employed for the regression analysis, and then we discuss our approaches to dealing with the serial correlation and heteroskedasticity issues. These discussions are followed by a listing of regression results obtained using four alternative estimation methods, and finally we summarize the technical orientation of the study.

A.I.1. The Regression Specification

The regression specification begins with a standard conditional demand form which is modified to take account of the Energy Use Index (EUI) information typically employed for commercial sector end-use modeling. For the standard conditional demand form, energy consumption is modeled as the sum of the loads for the end-use categories connected to the meter. In addition to this modeling principle there is generally some breakdown of loads by customer category. For example, in residential models we would frequently break down space-heating or air-conditioning loads into single-family or multi-family categories. For the commercial sector there is generally a similar breakdown of loads by building, or business, type.

To implement the building (business) - type breakdown of loads for the current study, the basic energy demand equation is specified to take account of the following principles:

1. metered energy consumption is sum of the loads for the end-use categories connected to the meter; and that,
2. end-use loads may vary by building-type as well as other typical determinants of energy consumption such as floorspace and weather.

To take account of the end-use and building-type breakdown, we can write the conditional demand energy use equation for a given building type as follows:

A.I.1.1) $$Thm = \sum_{i=1}^{EUcat} Thm_{i,bt} \cdot A_i$$

Table A.I.1
Definitions Of Terms In Equation A.I.1

Term	Units	Definition
Thm	Monthly *Thm*	*Thm* customer consumption per billing month
$Thm_{i,bt}$	Monthly *Thm*	*Thm* consumption per customer per billing month for end-use category *i* in building-type *bt*.
Eucat	end-use categories	The total number of end-use categories (we concentrate on 3 for the current study: space heating, cooking and 'other')
A_i	pure units	dummy variable for end-use category's presence: 1: end-use category *c* is connected to the meter, 0: end-use category *c* is not connected to the meter.

A common practice in energy forecasting, since Jerry Jackson's pioneering work in commercial end-use modeling, is to use prior information (generally, engineering priors) on energy use per square-foot along with end-use saturation information in the equation. Using this approach, we can rewrite Equation A.I.1 as follows:[36]

A.I.1.2) $$Thm = \sum_{i=1}^{EUcat} \left(\frac{Thm}{FS} \cdot S \right)_{i,bt} \cdot FS$$

Table A.I.2
Definitions Of New Terms In Equation Table A.I.2

Term	Units	Definition
$\left(\frac{Thm}{FS} \cdot S \right)_{i,bt}$	Monthly *Thm* per square foot	prior (or calibrated) estimate of *Thm* consumption per square foot of floorspace for end-use *i* in building-type *bt*, per billing month, weighted by the appliance saturation for end-use *i* in building-type *bt*
S	pure number	saturation for the relevant end-use category and building type combination; i.e., the expected value of the end-use category variable, given the building type -- in some references the end-use saturations are adjusted by the proportion of floor-space served by the end-use category (in terms of equation A.I.1.2, S = $\overline{A}$).
FS	square feet	square feet of floorspace in the area served by the meter

[36] In forecasting applications the *aggregate* Thm equation for building-type *bt* at the system level, Thm_{bt}, can be written:

$$Thm_{bt} = \sum_{c=1}^{EUcat} EUI_{c,bt} \cdot \overline{F}_{bt} \cdot Customers_{bt} = \sum_{c=1}^{EUcat} EUI_{c,bt} \cdot F_{bt}$$

where, $\overline{F}_{bt}$ is average floorspace for building-type *bt*, $Customers_{bt}$, and F_{bt} is the total system-level floorspace in building-type *bt*. In this type of setting the EUI's are generally at the annual level since annual movements in system-level Thm are of great interest for planning.

The $\left(\frac{Thm}{FS}\cdot S\right)_{i,bt}$ term can be considered a saturation-weighted EUI (Energy Use Index) for end-use *i*, in building type *bt*. The corresponding unweighted EUI can be written as $\left(\frac{Thm}{FS}\right)_{i,bt}$. Using this term, equation A.I.1.2 can be rewritten as

A.I.1.2a) $$Thm = \sum_{i=1}^{EUcat}\left(\frac{Thm}{FS}\right)_{i,bt}\cdot S_{i,bt}\cdot FS = \sum_{i=1}^{EUcat} EUI_{i,bt}\cdot S_{i,bt}\cdot FS$$

Note that in this equation, the EUI term is understood to be the unweighted EUI.[37] In our regression work we use the building-type saturations only where survey information about the presence of a particular type of equipment is missing. Where survey information exists, the $S_{i,bt}$'s are customer-specific dummy variables that indicate the presence or absence of equipment.

The end-use categories in our current study are space heating, cooking, and 'other'. Writing Equation A.I.2 explicitly for these categories we have (again, for a given building type, *bt*)

A.I.1.3) $$Thm = (EUI_{h,bt}\cdot S_{h,bt} + EUI_{k,bt}\cdot S_{h,bt} + EUI_{o,bt}\cdot S_{h,bt})\cdot FS$$

In this equation, the EUI indices *h*, *k*, and *o*, represent the HVAC, cooking and 'other' end-use categories, respectively. The FS_h variable represents the conditioned (heated and cooled) floor-space, and the *FS* variable represents the overall floor space.

Since we expect the loads for space heating to be weather-sensitive, we create interaction terms to adjust the heating EUI's for the effects of changing weather conditions. This weather variable is constructed as a ratio to normal weather so that they affect the EUI's only to the extent that current weather conditions differ from normal weather. In the remaining discussion then we will use the following equation when referring to the space heating EUI.

A.I.1.4) $$EUI_{sh,bt} = (nhdd_m\cdot EUI_{sh,bt})$$

[37] The EUI's and Saturations are typically presented in two forms: 1) taking the total floor space into account; and, 2) taking into account only the floor space served by end-use *i* in business-type bt. In appendix II we present *both* types of EUI's and saturations for the interested reader. In our regression analysis we use the EUI's and Saturations that are applied to the *total* floor space.

Table A.I.3
Definitions Of New Terms In Equation A.1.4

Term	Units	Definition
$EUI_{sh,b}$	monthly Thm /sqft	energy use index for space-heating in building-type *bt*
$Nhdd_m$	pure number	*recorded* monthly heating degree-days divided by *normalized* monthly heating degree-days, in month *m*

Of course the EUI's may not correspond exactly to reality, and in our regression equation, we estimate an overall realization coefficient for the EUI's for customer within a particular building type, *bt*, as follows:

A.I.1.5) $$Thm = \alpha_{bt} \cdot [EUI_{sh,bt} \cdot S_{sh} + EUI_{ck,bt} \cdot S_{ck} + EUI_{o,bt} \cdot S_o] \cdot FS + e$$

or

A.I.1.5a) $$Thm = \alpha_{bt} \cdot [\sum_{i=sh}^{o} EUI_{i,bt} \cdot S_i] \cdot FS + e$$

A more general expression that takes into account all the building types in our sample is as follows:

A.I.1.5b) $$Thm = \sum_{bt=1}^{BT} \left[\alpha_{bt} \cdot d_{bt} \sum_{i=sh}^{o} EUI_{i,bt} \cdot S_i \right] \cdot FS + e$$

In this equation, α_{bt} represents the overall EUI realization coefficient for building type bt, d_{bt} is a 0,1 dummy variable that indicates whether the customer is of type *bt*, and the *e* term represents an additive error term.

To take account of the possibility that the end-use categories should have different weights, we create an EUI term in which we divide the end-use categories into space heating, cooking, and other, and test for incremental weights for two of these three categories as follows:

A.I.1.6) $$EUIterm_{bt} = \left[1 \cdot \left(\sum_{i=sh}^{o} EUI_{i,bt} \cdot S_i \right) + a \cdot EUI_{sh,bt} S_{sh} + c \cdot EUI_{ck,bt} S_{ck} \right]$$

where

A.I.1.6a) $$a = a_{sh} + a_{sh,97} \cdot d_{97}$$

In this term, the parameters *a* and *c* would be 0 if the estimated weighting for the three end-use categories were equal. The a_{sh} and a_{sh} parameters are used to test for general and year 1997 incremental weights, respectively, for the space heating category. The *c* parameter is further broken down to test whether the overall incremental cooking weight changes with the month, or with customer changes in the volume of business, or other general changes during our post program year, 1997. To carry out these tests, we specify

c as a function rather than a single parameter, as follows:

$$\begin{aligned} c = {} & a_{ck} + (a_{ck,HighVolume} \cdot dHighVolumeMonth) \\ & + [a_{ck,97} + (a_{ck,97,IncreasedActivity} \cdot dActivityUp) \\ & + (a_{ck,97,decreasedActivity} \cdot dActivityDown)] \cdot d97 \end{aligned} \qquad \text{A.I.1.7)}$$

where the a_{ck} parameters are incremental cooking weights to be estimated, *dHighVolumeMonth* is a 0,1 dummy variable that is 1 if the survey respondent claims that the current billing month is a high volume period. The *dActivityUp* and *dActivityDown* variables are 0,1 dummy variables that have the value 1 if the customer reported in 1997 that the volume of activity moved up or down, respectively from the program-year level. The d97 variable is a 0,1 dummy variable that is 1 if the current billing month is in the post-program year, 1997

The building types considered in the SoCalGas forecasting model are summarized in appendix A.II.1. In our sample, however, we made a further categorization of customers depending upon the floorspace information available for them. There were three categories of floorspace information: 1) the floorspace specified by the survey respondents, when that was less than 10,000 square feet; 2) a 0,1 dummy variable that has the value 1 only if the floorspace was greater than 10,000 square feet; and, 3) a 0,1 dummy variable that has the value 1 if the floorspace information is entirely missing. To account for these three categories we estimate a different coefficient for each building-type/square-footage category. For example, we estimate different α_{bt} coefficients for restaurants with specified and missing floorspace. As noted in section III, for respondents with no specified floorspace, the estimated coefficients account for differences in floorspace as well as differences in the applicability of the prior EUIs.

In our regression model, then, the BT summation limit in equation A.I.1.5b takes into account all the original SoCalGas building types and the three categories of floorspace information. The complete list of these α_{bt} is included in table III.1, *Regression Results*, in section III. In addition, as noted in section III, we also tested the 1% upper and lower limits of the floorspace distribution for possible variations from the average EUIs reported by the SoCalGas forecasting department.

To take the into account the complete floorspace specification, we replace the FS variable above with the following expression,

$$Fsterm = FSspec + d_{miss} + d_{10k} \qquad \text{A.I.1.8)}$$

where, *FSspec* is the actual floorspace (in square feet) where it is specified, and *FSspec* is 0 otherwise; d_{miss} is a 0,1 dummy variable that is 1 if the floor space is not specified, and d_{miss} is 0 otherwise; and, d_{10k} is a 0,1 dummy variable that is 1 if the customer's floor space is greater than 10,000 square feet, and d_{10k} is 0 otherwise.

Apart from these building type and floorspace-category adjustments to the EUI's we tested to see whether there were any other general differences across program participation categories and years and modified our therm equation as follows:

A.I.1.9) $$Thm = AdjNpYr \cdot \sum_{bt=1}^{BT} \left[\alpha_{bt} \cdot d_{bt} EUIterm\right] \cdot FSterm + e$$

where,

A.I.1.10) $$AdjNpYr = 1 + \beta_{npart} \cdot d_{npart} + \left(\beta_{96} + \beta_{part,96} \cdot d_{part}\right) \cdot d_{96} + \beta_{97} \cdot d_{97}.$$

In this expression, the β's are parameters to be estimated, d_{npart} is a dummy variable if the customer *did not* participate in the SoCalGas 1996 program, and d_{npart} is 0, otherwise; , d_{part} is a dummy variable if the customer *did* participate in the SoCalGas 1996 program, and d_{part} is 0, otherwise; d_{96} is a dummy variable if the current billing month is in 1996, and d_{96} is 0, otherwise; and, d_{97} is a dummy variable if the current billing month is in 1996, and d_{97} is 0, otherwise.

To test for any further general additive adjustments to the SoCalGas forecasting EUI's apart from the interaction terms in equation A.I.1.9, we modified our therm equation as follows:

A.I.1.11) $$Thm = AdjNpYr \cdot \sum_{bt=1}^{BT} \left[\alpha_{bt} \cdot d_{bt} EUIterm\right] \cdot FSterm + EUIni + e$$

where

A.I.1.12) $$EUIni = \gamma_{sh} \cdot EUI_{sh,rest} \cdot S_{sh,rest} + \gamma_{nonSh} \cdot EUI_{nonSh,rest} \cdot S_{nonsh,rest} + \phi_{sh} \cdot dfEUI_{sh,bt} + \phi_{nonSh} \cdot dfEUI_{nonSh,bt}$$

and,

A.I.1.13) $$dfEUI_{i,bt} = \frac{EUI_{i,bt} \cdot S_{i,bt} - EUI_{i,rest} \cdot S_{i,rest}}{EUI_{i,rest} \cdot S_{i,rest}};\ i = sh,\ ck,\ o$$ [38]

In these last equations $EUI_{i,bt}$ and $S_{i,rest}$ represent the EUI and saturation of end-use i in building-type bt, and the γ and ϕ symbols are parameters to be estimated.

As noted in section III, these γ and ϕ parameters (and the *EUIni* term) were included to account for any general *additive* adjustment to the overall EUI's, apart from floor space

[38] Note that $dfEUI_{nonSh,bt} = dfEUI_{ck,bt} + dfEUI_{o,bt}$

and the other interaction terms. For this purpose, we divided the end-use equipment into space heating and all-other categories. This part of the analysis was carried out in two logical steps: 1. a step to see if a standard set of EUIs for restaurants would suffice to account for an adjustment to the EUIs apart from the other EUI interaction terms;[39] and, 2. a step to test whether proportionate differences between each observation's EUIs and the restaurant EUIs were useful adjustments to the EUIs after taking into account the other interactions (floorspace, etc.). It should be stressed that the inclusion of the EUIni term in the equation affects our interpretation of the α_{bt} parameters: these parameters are no longer interpreted as realization rates of the SoCalGas EUIs.

A final term was added to the equation to test whether there was any evidence that there were any program impacts after controlling for general movements in consumption across the years in our sample, and after controlling for size and activity-level differences between the program participants. These latter issues are addressed in conformity with the requirements of the California Protocols. The complete estimation equation can, therefore be written as follows:

A.I.1.14)
$$Thm = AdjNpYr \cdot \sum_{bt=1}^{BT} \left[\alpha_{bt} \cdot d_{bt} EUIterm\right] \cdot FSterm + EUIni + (r + \theta \cdot AddEquip) \cdot \left(M \cdot d_{part} \cdot d_{97}\right) + e$$

In this equation the M variable represents the per-installation *ex ante* net impact filed by SoCalGas for its 1996 equipment replacement program.[40] M is calculated by multiplying the Gross Impact per Unit (installation) by the *ex ante* Net-To-Gross ratio. M is therefore 749 ≈ (998 therms per year) * (.75, the *ex ante* Net-To-Gross ratio).[41] The negative of the parameter *r* is the expected value of the *net* impact per installation realization rate. If the negative of this parameter is greater than or equal to 1.0, then the parameter is consistent with the view that the program met or exceeded its *ex ante* estimates, otherwise the parameter would suggest that the program did not meet its *ex ante* estimates. We did not build a scale effect into this estimate since there seem to be significant determinants of cooking apart from overall floorspace or consumption. The $(\theta \cdot AddEquip)$ term is used to account for the effects of increases in participants' stock of cooking equipment. AddEquip is a 0,1 dummy variable that is 1 if the participant added, rather than replaced, any equipment in connection with the program.

Equation A.I.1.14 is our basic overall regression specification. To summarize the model, it includes coefficients designed to account for adjustments to the EUIs by building type, and end-use category, and floorspace information category. In addition, the model

[39] As noted above, the effect of adding the nonspace-heating portion of the restaurant EUI term is to add a constant term to the equation.

[40] See Appendix II (section A.II.2) for the source references for the SoCalGas *ex ante* filings.

[41] Note that since we perform a monthly analysis, we further divide this number (749) by 12. A similar adjustment is made to translate the annual EUI's obtained from the SoCalGas forecasting department to monthly values.

contains coefficients that account for general EUI differences across program participation categories and sample years (pre-program period, program year, post program period). Furthermore, the model contains coefficients that account for EUI effects apart from all the other interactions contained in the model. Finally, the model contains a term used to estimate the net impact using the California Protocol requirements for impact measurement. Although equation A.I.1.14 is written in a slightly compressed form, it has the same components as the overall estimating equation III.A.2 in section III of the text.

A.I.2. The Estimation Procedure

In this section, we describe the nonlinear regression procedure that we use to estimate the parameters of our model. After this we describe our approach to the serial correlation and heteroskedasticity issues and summarize the regression results obtained from four alternative estimation procedures.

A.I.2.a The Nonlinear Estimation Procedure

Since the difference equation specification is nonlinear, we employ a nonlinear estimation routine in which we obtain the parameters using an iterative estimation approach. The approach we use is called a *linearized regression model*[42] and is based upon a Taylor series approximation to a function. This procedure works as follows. Suppose that we write our regression as

A.I.2.a.1) $Thm = f(x,b) + e$

where *Thm* represents the monthly consumption of Therms, $f(x,b)$ represents our commercial energy demand equation, x represents our vector of explanatory variables, b represents our vector of parameters, and u is an additive error term. We can write a first-order, linear, Taylor series approximation to this function given a set of initial values of the elements of the b vector, b^0, as follows:

A.I.2.a.2) $$Thm = f(x,b^0) + \sum_{p=1}^{TotParms} \frac{\partial f(x,b^0)}{\partial b_p} \cdot (b_p - b_p^0) + e,$$

where, *TotParms* represents the number of parameters to be estimated.

Using the linearization approach, then, we can transform Equation A.I.2.2 as follows:

A.I.2.a.3) $$ThmAdj = \sum_{p=1}^{TotParms} \frac{\partial f(x,b^0)}{\partial b_p} \cdot (b_p) + e$$

where we define,

[42] See, for example, W.H. Greene, *Econometric Analysis*, Macmillan Publishing Co., New York, 2nd Ed., pp. 315-316.

A.I.2.a.4) $$ThmAdj \equiv Thm - f(x,b^0) + \sum_{p=1}^{TotParms} \frac{\partial f(x,b^0)}{\partial b_p} \cdot (b_p^0).$$

To implement this approach, we use A.I.2.a.3 as a template for a regression in which we employ *ThmAdj* as a dependent variable, and we use the derivatives on the right side of the equations as regressors. In order to compute *ThmAdj* and the right-side variables, we specify an initial set of parameter values, b^0, and we use these to compute an initial value of predicted *Thm* using $f(x,b^0)$, and the appropriate set of derivatives.

Given, these derivatives and the *ThmAdj* variable, the procedure described above is continued employing a Newton[43] approach in which updated estimates of the b vector are used at every step to compute the indicated function values, the $f(x,b^0)$ term, and the indicated derivatives of the function with respect to the parameters. This iterative reestimation process is continued until none of the parameters changes by more than 1% (in absolute value) from one iteration to the next.[44]

A.I.2.b. The Serial Correlation And Heteroskedasticity Issues

In this section, we will outline our approaches to the serial correlation and heteroskedasticity issues.

Serial Correlation

For Ordinary Least Squares (OLS) regression analysis, we typically assume that the current period's error is independent of errors in previous periods. If this is not true, we can attempt to build a model that takes this temporal error dependence into account. A typical case that is modeled is that of first-order error dependence, or serial correlation, into account. Although such dependence will not bias the OLS regression coefficients, greater efficiency will result if we take this information into account.

Applying the conventional approach to our current model consider our basic equation, modified to make the time period explicit:

A.I.2.b.1 $Thm_t = f(x_t,b) + e_t$.

The first-order serial correlation specification can be written as follows

A.I.2.b.2 $e_t = \rho \cdot e_{t-1} + u_t$.

For a linear equation it is customary to purge the serial correlation from the error term by replacing each variable in the model with a generalized difference form in which the *lagged* value of each variable is multiplied by an estimate of ρ. This product is then

[43] Again, Greene, op. cit., page 347, is a good reference.

[44] The parameter percentage changes from iteration to iteration are computed as $\left[\frac{param_{new} - param_{old}}{param_{old}}\right] \cdot 100 \cdot$

subtracted from the *current* value of the variable. Although the current model is nonlinear, we can perform a similar operation by rewriting our equation in the following generalized difference form.

A.I.2.b.3 $Thm_t - \rho \cdot Thm_{t-1} = f(x_t,b) - \rho \cdot f(x_{t-1},b) + u_t$.

Adding $\rho \cdot f(x_{t-1},b)$ to both sides of the equation this can be rewritten

A.I.2.b.4 $Thm_t - \rho \cdot [Thm_{t-1} - f(x_{t-1},b)] = f(x_t,b) + u_t$.

or

A.I.2.b.5 $Thm_t - \rho \cdot [e_{t-1}] = f(x_t,b) + u_t$.

We can obtain a consistent estimate of e_t using the equation

A.I.2.b.6 $\hat{e}_t = Thm_t - f(x_t, \hat{b})$.

This estimate can be used to obtain an estimate of ρ in the context of the following regression equation:

A.I.2.b.7 $\hat{e}_t = \rho \cdot \hat{e}_{t-1} + u_t$.[45]

Using the estimates of ρ and e_{t-1} we can rewrite equation A.I.2.b.6 as follows:

A.I.2.b.8 $Thm_t - \hat{\rho} \cdot \hat{e}_{t-1} = f(x_t,b) + u_t$.

This method is one of a number of straightforward ways to address the serial correlation of residuals for nonlinear regression functions. For the current case, however, a further issue arises because of the possibility that the *variance of* u_t, the portion of the error term purged of serial correlation, may not be constant across customers.

Heteroskedasticity

In conventional regression analysis, the error term is viewed as having a variance that may or may not be constant across the observations. In rough terms, this error variance may be regarded as an indicator of the noise that is attached to a particular observation. In standard econometric texts, the optimality of the regression estimates is shown to be conditional upon the assumption that the error variance is constant. In general, if the error variance is not constant, the remedy is to apply a set of weights to the data. The suggested weights are higher for the observations with higher error variances. The optimal weights are constructed in such a way that the resulting *weighted* error variances are equal across observations.[46]

[45] To simplify the notation we have not distinguished between the u_t in this equation and the corresponding variable in equation A.I.2.b.2.

[46] See, for example, Greene, op. cit.

Note that we can obtain an estimate of u_t by solving an estimate of equation A.I.2.b.8 as follows:

A.I.2.b.9 $\hat{u}_t = [Thm_t - f(x_t, \hat{b})] - \hat{\rho} \cdot \hat{e}_{t-1}$.

or

A.I.2.b.10 $\hat{u}_t = \hat{e}_t - \hat{\rho} \cdot \hat{e}_{t-1}$.

An estimate of σ_u^2, the variance of u_t, can be obtained by squaring equation *A.I.2.b.10.*

In the current study, we estimate an error variance regression equation that takes into account the floorspace, for survey respondents whom that information was available, a dummy variable that indicates whether the floor space information was available, a dummy variable indicating whether the floor space exceeded 10,000 square feet, and a dummy variable indicating the program participation status of the customer.

Dependent Variable: $\ln(\hat{\sigma}_u^2)$	
Definition Of Variable	**Regression Parameter Estimates (t-ratios)** ***generalized difference* regression** $R^2 = .87$ *Number Of Observations =28,889*
intercept	5.62 (132.32)
1 =>square feet known and less than 10,000 square feet 0 , otherwise	0.0002 (19.00)
dummy variable: 1 =>square feet unknown 0 , otherwise	1.28 (27.33)
dummy variable: 1 =>exact square feet unknown, except that it is greater than 10,000 0 , otherwise	5.91 (22.29)
dummy variable: 1 =>customer is a program participant 0 , otherwise	0.84 (25.44)

Denoting the estimated value of this equation $\ln S^2$, we computed the regression weight for each observation as follows:

A.I.2.b.10 RegressionWeight = $[\exp(\ln S^2)]^{-.5}$

Both sides of the regression equation are divided by this weight.

A.I.3. Regression Results For Alternative Estimation Procedures

In section III of this study we presented our regression results using a nonlinear estimation routine that addressed the standard serial correlation and heteroskedasticity issues that arise in analyses of pooled time-series/cross-section data sets. We noted, however, that we had also estimated the model using several other approaches. Since our model is nonlinear, we employed the nonlinear methods described above for each of these approaches.

In this section we report the results of all the alternative analyses. As a convenience for the reader we first reproduce the equation specification from section III since we will refer to the symbols in these equations again in this section. The definitions of the symbols will be given in context.

$$\begin{aligned} Thm = {} & (AdjNpYr)\cdot\sum_{bt=1}^{BT}\alpha_{bt}\cdot d_{bt}\cdot EUIFS_{bt} \\ & +\gamma_{sh}\cdot EUI_{sh,rest}+\gamma_{nonSh}\cdot EUI_{nonSh,rest} \\ & +\phi_{sh}\cdot dfEUI_{sh,bt}+\phi_{nonSh}\cdot dfEUI_{nonSh,bt} \\ & +(r+\theta\cdot AddEquip)\cdot(M\cdot d_{part}\cdot d_{97})+e. \end{aligned} \qquad \text{III.A.2)}$$

where,

$$AdjNpYr = 1+\beta_{npart}\cdot d_{npart}+\left(\beta_{96}+\beta_{part,96}\cdot d_{part}\right)\cdot d_{96}+\beta_{97}\cdot d_{97} \qquad \text{III.A.3)}$$

$$EUIFS_{bt} = \left[1\cdot\left(\sum_{i=sh}^{o} EUI_{i,bt}\cdot S_i\right)+a\cdot EUI_{sh,bt}S_{sh}+c\cdot EUI_{ck,bt}S_{ck}\right]\cdot FSterm \qquad \text{III.A.4)}$$

$$a = a_{sh}+a_{sh,97}\cdot d_{97} \qquad \text{III.A.4a)}$$

$$\begin{aligned} c = {} & a_{ck}+(a_{ck,HighVolume}\cdot dHighVolumeMonth) \\ & +[a_{ck,97}+(a_{ck,97,IncreasedActivity}\cdot dActivityUp) \\ & +(a_{ck,97,decreasedActivity}\cdot dActivityDown)]\cdot d97 \end{aligned} \qquad \text{III.A.5)}$$

$$FSterm = FSspec + d_{miss} + d_{10k} \qquad \text{III.A.6)}$$

III.A.7) $dfEUI_{i,bt} = \dfrac{EUI_{i,bt} \cdot S_{i,bt} - EUI_{i,rest} \cdot S_{i,rest}}{EUI_{i,rest} \cdot S_{i,rest}}$; $i = sh, ck, o$ [47]

The methods we used can be summarized as follows:

Procedure	coefficient, ρ, of lagged value in difference term[48]	Regression Weight
1. weighted, *generalized* first difference	.9528 (estimated value of the serial correlation coefficient ρ)	estimated from relevant $\ln(\hat{\sigma}_u^2)$ equation
2. first-difference	1	1
3. ordinary least squares	0	1

In these regressions, the weighted, generalized difference regressions gave estimated savings impacts that were similar to the SoCalGas *ex ante* estimates. The first-difference and ordinary least squares impact estimates were more than double that of the SoCalGas *ex ante* estimates.

For the α_{bt} parameters that adjust for EUI's across business-type and types of floorspace information, the Ordinary Least Squares results are generally similar in magnitude to the weighted generalized difference parameters. By contrast, the first difference regression has α_{bt} estimates that are generally higher than those for ordinary least squares or the weighted, generalized difference regression. Recall, however, that the first-difference regression deletes two EUI adjustment parameters, γ_{nonSh} and ϕ_{nonSh}, contained in the other regressions.[49]

We show the results obtained from each of these procedures in the following table.

[47] Note that $dfEUI_{nonSh,bt} = dfEUI_{ck,bt} + dfEUI_{o,bt}$

[48] We are referring here to the value of $\hat{\rho}$ in equation A.I.2.b.8. This equation was written:

$$Thm_t - \hat{\rho} \cdot \hat{e}_{t-1} = f(x_t, b) + u_t .$$

[49] The terms containing these parameters are constant for an individual customer. Therefore they cancel to zero in our first-difference regressions.

Table A.1.3
Regression Result Summary

Dependent Variable: Standardized Therms Per Customer Per Month (*Thm*)				
Symbol In Equations III.A.2 - III.A.7	**Symbol Definition**	**Regression Parameter Estimates (t-ratios)**		
		1 *Weighted, Generalized Difference*	2 *First Difference*	3 *Ordinary Least Squares*
r	realization coefficient for *ex ante* estimates of net impacts associated with the 1996 SoCalGas ER DSM program	-1.06 (-8.23)	-3.17 (-19.43)	-2.10 (-4.41)
θ	increment to net impact realization coefficient to account for additional, rather than replacement, equipment installed in connection with program.	0.33 (1.44)	1.23 (5.10)	2.16 (2.77)
β_{npart}	overall difference (roughly proportionate) in consumption between nonparticipants and participants.	-0.48 (-106.09)	-0.50 (-151.58)	-0.57 (-33.96)
β_{96}	overall change (roughly proportionate) in consumption from the pre-program period to the program year, 1996.	-0.04 (-6.90)	-0.04 (-7.23)	0.00 (0.07)
$\beta_{part,96}$	increment to overall change in consumption from the pre-program period to 1996 for participants. This may result from program-related effects during the program year. These effects are not, however, included in our calculation of pre-post impact differences.	-0.02 (-2.41)	-0.09 (-17.74)	0.02 (0.66)
β_{97}	overall change (roughly proportionate) in consumption from the pre-program period to the post-program year 1997.	-0.12 (-17.07)	-0.17 (-61.54)	-0.03 (-1.59)
$\alpha_{restaurant}$	overall adjusted EUI coefficient for the base year (1995): restaurant category	0.63 (21.81)	3.30 (75.78)	0.69 (7.59)
$\alpha_{foodStore}$	overall adjusted EUI coefficient for the base year (1995): food store category	3.19 (22.29)	10.60 (41.74)	3.47 (7.72)
$\alpha_{K\text{-}12\ schools}$	overall adjusted EUI coefficient for the base year (1995): K-12 school category	25.19 (51.13)	72.88 (109.09)	23.66 (20.06)

Dependent Variable: Standardized Therms Per Customer Per Month (*Thm*)				
Symbol In Equations III.A.2 - III.A.7	**Symbol Definition**	**Regression Parameter Estimates (t-ratios)**		
		1 *Weighted, Generalized Difference*	2 *First Difference*	3 *Ordinary Least Squares*
$\alpha_{healthFacilities}$	overall adjusted EUI coefficient for the base year (1995): health-facilities category	3.42 (40.54)	7.34 (62.42)	2.18 (6.37)
$\alpha_{miscellaneous}$	overall adjusted EUI coefficient for the base year (1995): miscellaneous building-type category	3.98 (23.56)	27.75 (131.15)	0.82 (1.94)
$\alpha_{FS\ upper\ 1\%}$	increment to EUIs for customers in upper 1% of floorspace distribution	-0.47 (-13.54)	-1.61 (-20.63)	-0.72 (-6.44)
$\alpha_{FS\ lower\ 1\%}$	increment to EUIs for customers in lower 1% of floorspace distribution	3.93 (17.05)	14.27 (11.67)	5.24 (3.72)
$\alpha_{restaurant,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): restaurant category, observations with missing floorspace	2313.74 (21.72)	12059.50 (78.51)	2630.05 (7.53)
$\alpha_{K\text{-}12\ schools,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): K-12 school category, observations with missing floorspace	78955.32 (54.23)	211233.08 (116.65)	88518.43 (24.76)
$\alpha_{healthFacilities,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): health-facilities category, observations with missing floorspace	35525.71 (130.21)	60827.50 (266.43)	35601.78 (51.70)
$\alpha_{hotel/motel,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): hotel/motel category, observations with missing floorspace	132629.68 (39.34)	260985.81 (162.80)	134428.43 (14.93)
$\alpha_{miscellaneous,\ missing\ floorspace}$	overall adjusted EUI coefficient for the base year (1995): miscellaneous building-type category, observations with missing floorspace	49640.64 (43.78)	218937.35 (174.29)	39050.81 (14.02)
$\alpha_{colleges,\ floorspace>10,000\ sqft}$	overall adjusted EUI coefficient for the base year (1995): college category, observations with floorspace greater than 10,000 square feet	173210.74 (5.70)	487251.56 (95.30)	156295.35 (16.67)
$\alpha_{healthFacilities,\ floorspace>10,000\ sqft}$	overall adjusted EUI coefficient for the base year (1995): health-facilities category, observations with floorspace greater than 10,000 square feet	385304.17 (45.25)	623999.14 (510.18)	354956.78 (79.15)

Dependent Variable: Standardized Therms Per Customer Per Month (*Thm*)				
Symbol In Equations III.A.2 - III.A.7	**Symbol Definition**	**Regression Parameter Estimates (t-ratios)**		
		1 *Weighted, Generalized Difference*	2 *First Difference*	3 *Ordinary Least Squares*
$\alpha_{hotel/motel,\ floorspace>10,000\ sqft}$	overall adjusted EUI coefficient for the base year (1995): hotel/motel category, observations with floorspace greater than 10,000 square feet	162051.07 (11.23)	217573.07 (113.43)	172227.70 (18.07)
a_{sh}	incremental EUI coefficients for the space heating end-use category.	-0.37 (-25.85)	-0.77 (-213.68)	-0.36 (-13.02)
$a_{sh,97}$	incremental EUI coefficients for the space heating end-use category in the post period, 1997.	0.09 (4.87)	0.05 (9.07)	-0.03 (-0.68)
a_{ck}	incremental EUI coefficients for the cooking end-use category.	1.11 (7.49)	-1.38 (-135.80)	0.48 (1.37)
$a_{ck,97}$	incremental EUI coefficients for the cooking end-use category in 1997.	0.59 (7.28)	0.34 (20.05)	0.03 (0.14)
$a_{ck,97,\ increased\ activity}$	incremental EUI coefficients for the cooking end-use category in 1997 for customers who claimed increased volume of activity in their establishments.	0.33 (5.07)	0.05 (3.36)	0.01 (0.04)
$a_{ck,97,\ decreased\ activity}$	incremental EUI coefficients for the cooking end-use category in 1997 for customers who claimed decreased volume of activity in their establishments.	-0.32 (-4.23)	-0.05 (-2.95)	-0.55 (-2.60)
$a_{ck,HighVolume}$	incremental EUI coefficients for the cooking end-use category in the high volume months identified by respondents.	0.38 (6.75)	0.03 (2.64)	1.85 (6.89)

Dependent Variable: Standardized Therms Per Customer Per Month (*Thm*)				
Symbol In Equations III.A.2 - III.A.7	**Symbol Definition**	**Regression Parameter Estimates (t-ratios)**		
		1 *Weighted, Generalized Difference*	2 *First Difference*	3 *Ordinary Least Squares*
γ_{sh}	general *additive* adjustment to the space heating EUI, apart from floor space and the other interaction terms (the restaurant EUI is used for this general purpose)	3014.27 (23.23)	3570.38 (21.50)	3023.65 (5.24)
γ_{nonSh}	general *additive* adjustment to the nonspace heating EUI, apart from floor space and the other interaction terms (the restaurant EUI's are used for this general purpose)	3331.60 (161.49)	---[50]	3479.01 (34.98)
ϕ_{sh}	incremental *additive* adjustment to the space heating EUI accounting for differences between the customer's business-type EUI for space heating and the restaurant EUI for space heating, apart from floor space and the other interaction terms	-14.11 (-12.17)	-20.39 (-12.89)	-22.79 (-4.42)
ϕ_{nonSh}	incremental *additive* adjustment to the nonspace heating EUI accounting for differences between the customer's business-type EUI for nonspace heating end-uses and the restaurant EUIs for the nonspace heating end-uses, apart from floor space and the other interaction terms	20.00 (6.87)	---	21.07 (1.69)

[50] These terms are constant for each customer, and the first difference would therefore be zero.

A.I.4. Technical Issues: Summary

The basic form of the regression equation that underlies our impact analysis is based upon the EUI framework that is commonly used in commercial forecasting models. We address the standard econometric issues in a straightforward way. To address the serial correlation and heteroskedasticity issues, we estimated a weighted, generalized difference regression. We also estimated an equation using first differences as an alternative to the *generalized* difference form.[51] The first difference form is interesting since serial correlation issues are typically avoided with such specifications. In addition, the self-selection bias issue is somewhat simplified for this specification since the inverse Mills ratio term cancels when the first-difference form is used.[52]

Finally, the basic form of our regression equation is nonlinear. To estimate the parameters of this equation, we employ a standard iterative approach called the *linearization* method. This method proceeds from a Taylor-Series linearization of the function to be estimated.

---

[51] Note, in the generalized difference form, an estimated serial correlation coefficient is used to weight the lagged value when computing the difference between the current and lagged value. By contrast, for the first-difference regression, the weight given the lagged value is 1.

[52] Note that the self-selection issue is concerned with the characteristics of the customer, not the consumption time period. Thus the Mills ratio term is calculated from customer-level cross-section variables. These variables, like any cross-section variable, cancel when we calculate period-to-period differences in consumption.

Appendix II: Prior Energy Use Indices And Savings Estimates

In this appendix we present the prior energy use indices and the prior savings estimates which were used in our analysis.

A.II.1 Prior Energy Use Indices And Floor-Space

Table A.II.1.a presents the prior energy use indices, in annual kBTU/SF, for commercial building types. These are estimates of 1991 SoCalGas EUI's which were prepared by Southern California Gas Co. in 1992.[53] The 'Other' category here is the sum of the remaining EUI's that are unspecified in this table.

Table A.II.1.a
Prior Energy Use Indices (EUI's)
Annual kBTU/SF

Building Type	Heating	Cooling	Water Heat	Cooking	Drying	Other
1. Office	9.78	0.00	5.92	**5.14**	0.00	0.00
2. Restaurant	21.80	0.00	96.93	**119.96**	0.00	0.71
3. Retail	9.05	0.00	12.93	**25.69**	0.00	3.62
4. Food Store	4.31	0.00	16.53	**41.39**	0.00	0.00
5. Warehouse	0.48	0.15	3.92	**0.00**	0.00	5.37
6. K - 12 Schools	20.63	5.28	8.36	**1.89**	0.00	3.38
7. Colleges & Universities	14.28	4.62	10.02	**26.83**	0.00	24.13
8. Health Facilities	27.24	40.24	74.10	**4.16**	0.00	2.02
9. Hotels & Motels	15.07	36.22	20.78	**15.26**	0.00	15.07
10. Laundries	114.80	0.00	73.97	**0.00**	20.60	26.18
11. Miscellaneous	8.28	0.00	24.35	**13.97**	0.00	7.39

[53] The data in the following table are taken from pages X-7 to X-17 in the SoCalGas, "Analysis of the 1992 Southern California Gas Commercial Survey"

Table A.II.1.b presents the equipment saturations for commercial building types based upon the commercial floor space *served by natural gas.* These are estimates of 1992 SoCalGas equipment saturations.[54]

Table A.II.1.b
Commercial Equipment Saturations And *Gas-Served* Floor Space

	Proportion of *Gas-Served* Floor Space						
Building Type	**Heating**	**Cooling**	**Water Heat**	**Cooking**	**Drying**	**Other**	Total *Gas-Served* Floor Space (Msqft)
1. Office	.8628	.0000	.7883	**.0845**	.0000	1.0000	567.70
2. Restaurant	.6314	.0000	.9597	**.8423**	.0000	1.0000	81.25
3. Retail	.5381	.0000	.2361	**.1859**	.0000	1.0000	122.94
4. Food Store	.5119	.0000	.5251	**.3133**	.0000	1.0000	46.99
5. Warehouse	.3819	.0142	.7233	**.0000**	.0000	1.0000	242.18
6. K - 12 Schools	.8066	.0580	.7813	**.3255**	.0000	1.0000	196.27
7. Colleges & Universities	.8481	.0004	.9344	**.0408**	.0000	1.0000	126.51
8. Health Facilities	.9541	.0452	.8719	**.3948**	.0000	1.0000	124.33
9. Hotels & Motels	.5219	.0078	.9525	**.4690**	.0000	1.0000	122.19
10. Laundries	.0353	.0000	.9715	**.0000**	.8373	1.0000	64.69
11. Miscellaneous	.6501	.0000	.8172	**.3127**	.0000	1.0000	263.75

[54] The data in the following table are taken from pages X-7 to X-17 in the SoCalGas, "Analysis of the 1992 Southern California Gas Commercial Survey"

Table A.II.1.c presents the equipment saturations for commercial building types based upon the *total* commercial floor space. These are estimates of 1992 SoCalGas equipment saturations.[55]

Table A.II.1.c
Commercial Equipment Saturations And *Total* Floor Space

Building Type	Proportion of *Total* Floor Space						
	Heating	Cooling	Water Heat	Cooking	Drying	Other	Total Floor Space (Msqft)
1. Office	.7275	.0000	.6647	**.0712**	.0000	1.0000	673.30
2. Restaurant	.5169	.0000	.7857	**.6896**	.0000	1.0000	99.30
3. Retail	.2487	.0000	.1007	**.0793**	.0000	1.0000	288.20
4. Food Store	.3337	.0000	.3423	**.2042**	.0000	1.0000	72.10
5. Warehouse	.2312	.0086	.4378	**.0000**	.0000	1.0000	401.10
6. K - 12 Schools	.7889	.05676	.7642	**.3184**	.0000	1.0000	196.27
7. Colleges & Universities	.6044	.0003	.6659	**.0291**	.0000	1.0000	177.50
8. Health Facilities	.7875	.0373	.7197	**.3259**	.0000	1.0000	150.60
9. Hotels & Motels	.4700	.0070	.8577	**.4223**	.0000	1.0000	135.70
10. Laundries	.0353	.0000	.9715	**.0000**	.8373	1.0000	64.70
11. Miscellaneous	.5182	.0000	.6515	**.2493**	.0000	1.0000	330.80

[55] The data in the following table are taken from pages X-7 to X-17 in the SoCalGas, "Analysis of the 1992 Southern California Gas Commercial Survey"

A.II.2. Ex-Ante (Prior) Gross Savings Estimates

Columns A and B of Table A.II.2 below contain the *ex-ante* program savings estimates that were used to construct priors, adjusted by realization coefficients, in our regression equation. The ex-ante impact estimates are on a per-project basis and are shown broken down by program and end-use category. *Note that although we present all of the savings categories of the 1996 SoCalGas commercial DSM programs, we only investigated the cooking end-use in our study.*

The *ex-ante* gross impacts are shown in column A, and the corresponding *ex-ante* net-to-gross ratios are shown in column B. These net-to-gross ratios are taken from the 1996 SoCalGas program goals filing (Advice Letter Supplement, 1995).

Table A.II.2
Ex Ante Program Impact Estimates For The SoCalGas 1996 Commercial Equipment Replacement Program[56]
(All Impact Estimates In Annual Therms Per Project)

End-Use	_Gross_ Impact Per Unit (A)	Net-To-Gross Ratio (B)	_Net_ Impact Per Unit (A x B)
Cooking[57]	998	.75	749
Water Heating[58]	498	.75	374
Boilers[59]	3,363	.75	2522
Gas Engines[60]	11,100	.75	8325
R-19 Ceiling Insulation[61]	692	.75	519
Air-Conditioning (Double for Single)[62]	64,982	.75	48,737
Air-Conditioning (Double Effect)[63]	-67,691	.75	50,768

The prior net savings estimates used in our analysis, the M_i in equation III.A.2, were constructed by multiplying these net-to-gross impact ratios by the *ex-ante* gross impact estimates presented in table A.II.2. We present these ex-ante net savings estimates in column C of this table.

---

[56] This information is taken from the, "*Southern California Gas Company Demand-Side management Advice Letter Attachments, 1996 Programs,*" October 1, 1995.
[57] See page 5.A.29, SoCalGas, op. cit.
[58] See page 5.A.25, SoCalGas, op. cit.
[59] See page 5.A.27, SoCalGas, op. cit.
[60] See page 5.A.31, SoCalGas, op. cit.
[61] See page 5.A.33, SoCalGas, op. cit.
[62] See page 5.A.35, SoCalGas, op. cit.
[63] See page 5.A.37, SoCalGas, op. cit.

Appendix III: The Analytic Data Base

In this appendix we give a detailed description of the data base which was used to support the statistical analysis. The first section presents the definitions of all variables included in the final data base. The second section lists the summary statistics for the included variables: the mean, range and number of valid and missing observations.

A.III.1. The Data Dictionary

The variables which were included in the final analytic data base are shown in the table below. Billing and weather variables are time-series vectors; the vector length is 45 periods in every case. Where fewer billing periods are present, the vectors are padded with zeroes at the end. Survey variables are cross-sectional in nature. The survey variables are primarily coded so that 1=yes/true, 0=no/false, and -1=not applicable/not answered. Exceptions to this general rule are noted where appropriate.

Table A.III.1
Variables Included in the Analytic Data Base

VARIABLE	DEFINITION
CSID	Unique case ID
PStatus	Participation status. 1=participant, 2=non-participant
SICCode	SIC code (4-digit)
WZone	Weather zone
Billing & Weather Data	
TotPeriods	Total # billing periods of consumption
Therms(n)	Therm consumption for each billing interval, normalized to uniform 30.4-day month. (45-element vector)
BillDays(n)	Number of days in each billing interval. (45-element vector)
JulMon94(n)	Month number of each meter-read date, numbered so that 1 represents Jan '94, 2=Feb '94, ..., 13=Jan '95, ... , 25=Jan '96, ... , 48=Dec '97. *(The earliest meter read date is July '94 (month #7); the most recent, October '97 (month #46).)* (45-element vector)
Hdd65N(n)	Cumulative heating degree-days, specific to the appropriate weather zone, for each billing interval, normalized to a uniform 30.4-day month.
Survey Data	
SQFT	Square-footage occupied by this facility
YrBuilt	Calendar year of construction of this building
RebEquip	Some Gas Company Rebate Program equipment was installed at this location.
dReplaProg	Program equipment was purchased to replace old equipment
dAdditProg	Program equipment was purchased in addition to old equipment.
OldPGGas	Program equipment replaced old gas equipment.
OldPGEle	Program equipment replaced old electric equipment.
Surprise	Non-participant indicates unexpected participation in Gas Company rebate program.
NewCook	Customer has installed some other (i.e., other than program-rebated) cooking equipment since January 1, 1996.
dRepla	This "other" new equipment replaced the old equipment.
dAddit	This "other" new equipment was in addition to the old equipment.

VARIABLE	DEFINITION
OldGas	"Other" equipment replaced old gas equipment.
OldEle	"Other" equipment replaced old electric equipment.
GASINCR96	Customer has increased the amount of gas-fueled cooking equipment since 1/1/96.
GASDECR96	Customer has decreased the amount of gas-fueled cooking equipment since 1/1/96.
GASPCTCHG	Customer's estimate of percent change of above increase or decrease.
SRVBREAK	This facility serves breakfast.
SRVLUNCH	This facility serves lunch.
SRVDINNER	This facility serves dinner.
SRVNOMEALS	This facility serves no meals.
SRVOTHER	This facility serves meals other than breakfast, lunch or dinner.
PCTBREAK	Estimated percent of meals cooked which are for breakfast.
PCTLUNCH	Estimated percent of meals cooked which are for lunch.
PCTDINNER	Estimated percent of meals cooked which are for dinner.
NUMMEALS	Estimated number of meals served.
MEALSMORE	Number of meals served has increased since 1/1/96.
MEALSLESS	Number of meals served has decreased since 1/1/96.
NMEALS95	Estimated number of meals cooked per week suring the year 1995.
MTFAST	Type of meals served here: fast food (burgers, fries, tacos, etc.).
MTPIZZA	Type of meals served here: pizza or sandwiches baked on premise.
MTSELF	Type of meals served here: self-service cafeteria.
MTTABLE	Type of meals served here: table service.
MTOTHER	Type of meals served here: other.
DAYSPERWK	Number of days per week the kitchen facility operates.
YEARROUND	Kitchen facility operates year-round.
CLOSE01	Kitchen facility is closed during Jan.
CLOSE02	Kitchen facility is closed during Feb.
CLOSE03	Kitchen facility is closed during Mar.
CLOSE04	Kitchen facility is closed during Apr.
CLOSE05	Kitchen facility is closed during May.
CLOSE06	Kitchen facility is closed during Jun.
CLOSE07	Kitchen facility is closed during Jul.
CLOSE08	Kitchen facility is closed during Aug.
CLOSE09	Kitchen facility is closed during Sep.
CLOSE10	Kitchen facility is closed during Oct.
CLOSE11	Kitchen facility is closed during Nov.
CLOSE12	Kitchen facility is closed during Dec.
SAMELEVEL	Amount of cooking done in this kitchen facility is about the same throughout the year.
HIVOL01	Heavier volume experienced during Jan.
HIVOL02	Heavier volume experienced during Feb.
HIVOL03	Heavier volume experienced during Mar.
HIVOL04	Heavier volume experienced during Apr.
HIVOL05	Heavier volume experienced during May.
HIVOL06	Heavier volume experienced during Jun.
HIVOL07	Heavier volume experienced during Jul.
HIVOL08	Heavier volume experienced during Aug.
HIVOL09	Heavier volume experienced during Sep.
HIVOL10	Heavier volume experienced during Oct.
HIVOL11	Heavier volume experienced during Nov.
HIVOL12	Heavier volume experienced during Dec.
BCMORE	Customer has expanded into new building space since 1/1/95.
BCLESS	Customer has reduced or eliminated some building space since 1/1/95.
BCRENOV	Customer has renovated some building space since 1/1/95.
BCSAME	Cusomer has made no change in amount of building space used since 1/1/95.
BCMORESF	Estimated added square-footage as a result of expansion
BCLESSSF	Estimated decreased square-footage as a result of reduction
COOLFUEL	Fuel used by cooling system: 1=gas; e=electric

VARIABLE	DEFINITION
HEATPCT	Percent of total building area which is heated
HEATGAS	Heating system is gas.
HEATNEW95	New heating system installed since 1/1/95.
HEATREPLA	New heating system replaced old heating system.
HEATWASGAS	Old heating system was gas.
HEATNEWYR	Year of new heating system installation
HEATNEWMO	Month of new heating system installation
HEATNEWPCT	Estimated percentage of total building area which is affected by new heating system
WHEATGAS	Fuel used by current water heater(s) is gas.
WHEATNUM	Number of water heaters present at facility.
WHEATCAP	Total capacity (in gallons) of water heater(s)
WHEATNEW95	New water heater has been installed since 1/1/95.
WHEATWASGA	Old water heater was gas.
WHEATNEWYR	Year of installation of new water heater
WHEATNEWMO	Month of installation of new water heater
WHEATREPLA	New water heater replaced the old.
WHEATPCT	Percent of hot water heated by the new water heating system
DIWASHCHG	Some change in amount of dishwashing has occurred since 1/1/96
DIWASHMORE	Amount of dishwashing has increased since 1/1/96.
DIWASHLESS	Amount of dishwashing has decreased since 1/1/96.
OTHERGAS	Natural gas is used at this facility for some purpose other than cooking, water heating, space heating or cooling.
Sample Design Variables	
SampCell	Sample cell for participant weighting: 1 = non-restaurant, small 2 = non-restaurant, medium 3 = non-restaurant, large 11 = restaurant, small 12 = restaurant, medium 13 = restaurant, large
AnalyWgt	Analysis weight. Used to weight data base sample participants to represent population of participants.

A.III.2 Summary Statistics

Table A.III.2 below presents the summary statistics for all variables included in the analytic data base. The mean, minimum, and maximum values for each variable are shown, along with a count of the number of good and missing observations.

Table A.III.2
Summary Statistics for Variables Included in the Analytic Data Base

Variable	mean	minimum	maximum	# good obs	# bad obs	missing obs
CSID	17923.1976	10001.0000	29093.0000	825	0	0
SICCODE	5950.3767	2099.0000	8661.0000	823	0	2
SQFT	1410.2655	99.0000	9999.0000	825	0	0
WZONE	5.0036	1.0000	6.0000	825	0	0
PSTATUS	1.5794	1.0000	2.0000	825	0	0
TOTPERIODS	37.1164	14.0000	45.0000	825	0	0
THERMS- 1	803.8302	3.0400	21780.5900	825	0	0
THERMS- 2	852.1157	2.0966	22509.6300	825	0	0
THERMS- 3	967.1573	0.9500	30232.8000	825	0	0
THERMS- 4	1054.1762	1.0133	32158.9800	825	0	0
THERMS- 5	1042.8415	1.9613	33008.3200	825	0	0
THERMS- 6	1008.6612	3.0400	27552.8800	825	0	0
THERMS- 7	956.9629	2.8500	27297.1000	825	0	0
THERMS- 8	932.0627	3.1448	27203.1000	825	0	0
THERMS- 9	907.1994	4.1931	26858.9200	825	0	0
THERMS- 10	880.0220	3.1448	26216.9600	825	0	0
THERMS- 11	834.5769	2.8500	23205.6500	825	0	0
THERMS- 12	823.2322	1.0133	21638.5100	825	0	0
THERMS- 13	830.0528	1.0133	21858.6500	824	0	1
THERMS- 14	820.7277	1.0133	24542.2300	824	0	1
THERMS- 15	886.3685	3.2571	26400.8300	822	0	3
THERMS- 16	925.6933	1.7882	28683.7800	820	0	5
THERMS- 17	962.8976	2.1714	30826.4700	819	0	6
THERMS- 18	975.8021	3.9226	30848.9100	819	0	6
THERMS- 19	958.5403	2.8500	31627.7700	818	0	7
THERMS- 20	905.4969	3.1448	27764.6300	818	0	7
THERMS- 21	869.0490	2.5333	26146.0300	813	0	12
THERMS- 22	851.8762	1.0133	25094.7100	812	0	13
THERMS- 23	830.7308	3.0400	24263.3900	804	0	21
THERMS- 24	812.9430	3.0400	25097.2300	803	0	22
THERMS- 25	811.3054	2.8500	26701.6500	800	0	25
THERMS- 26	836.6185	2.0966	27667.1400	796	0	29
THERMS- 27	885.1796	3.6848	27364.9000	793	0	32
THERMS- 28	925.9745	3.0400	29944.8700	784	0	41
THERMS- 29	946.0472	3.5765	28384.7700	775	0	50
THERMS- 30	949.3096	3.6848	26670.2300	771	0	54
THERMS- 31	916.5042	3.1448	28403.0300	766	0	59

THERMS- 32	858.9677	3.1448	22161.6000	755	0	70
THERMS- 33	826.2751	3.1448	24632.1100	744	0	81
THERMS- 34	809.1377	3.0400	23112.3900	734	0	91
THERMS- 35	801.8277	3.9226	20272.0500	723	0	102
THERMS- 36	831.0326	2.0267	26953.6500	705	0	120
THERMS- 37	776.3284	3.2571	10081.2700	641	0	184
THERMS- 38	794.6248	3.2571	8783.5030	433	0	392
THERMS- 39	762.7388	48.6400	5897.6000	46	0	779
THERMS- 40	1445.9587	95.1652	6674.9710	6	0	819
THERMS- 41	60.8000	60.8000	60.8000	1	0	824
THERMS- 42	0.0000	0.0000	0.0000	0	0	825
THERMS- 43	0.0000	0.0000	0.0000	0	0	825
THERMS- 44	0.0000	0.0000	0.0000	0	0	825
THERMS- 45	0.0000	0.0000	0.0000	0	0	825
BDAYS- 1	29.5067	4.0000	61.0000	825	0	0
BDAYS- 2	29.8824	11.0000	49.0000	825	0	0
BDAYS- 3	30.7661	10.0000	49.0000	825	0	0
BDAYS- 4	31.2521	9.0000	49.0000	825	0	0
BDAYS- 5	32.2206	14.0000	50.0000	825	0	0
BDAYS- 6	30.2933	10.0000	50.0000	825	0	0
BDAYS- 7	30.1503	8.0000	44.0000	825	0	0
BDAYS- 8	29.6206	13.0000	44.0000	825	0	0
BDAYS- 9	29.5224	13.0000	46.0000	825	0	0
BDAYS- 10	30.6291	12.0000	49.0000	825	0	0
BDAYS- 11	30.3321	6.0000	48.0000	825	0	0
BDAYS- 12	29.7491	13.0000	41.0000	825	0	0
BDAYS- 13	30.6311	6.0000	59.0000	824	0	1
BDAYS- 14	31.4867	2.0000	61.0000	824	0	1
BDAYS- 15	29.1180	15.0000	61.0000	822	0	3
BDAYS- 16	32.0146	5.0000	67.0000	820	0	5
BDAYS- 17	32.8400	4.0000	36.0000	819	0	6
BDAYS- 18	30.5250	7.0000	59.0000	819	0	6
BDAYS- 19	30.1222	5.0000	35.0000	818	0	7
BDAYS- 20	29.3655	2.0000	48.0000	818	0	7
BDAYS- 21	29.7626	2.0000	59.0000	813	0	12
BDAYS- 22	30.3966	4.0000	35.0000	812	0	13
BDAYS- 23	30.2550	6.0000	61.0000	804	0	21
BDAYS- 24	29.6849	1.0000	59.0000	803	0	22
BDAYS- 25	30.1563	4.0000	35.0000	800	0	25
BDAYS- 26	29.6382	3.0000	38.0000	796	0	29
BDAYS- 27	30.1475	4.0000	36.0000	793	0	32
BDAYS- 28	32.2717	6.0000	43.0000	784	0	41
BDAYS- 29	32.0981	6.0000	36.0000	775	0	50
BDAYS- 30	29.9546	4.0000	56.0000	771	0	54
BDAYS- 31	29.5235	3.0000	59.0000	766	0	59
BDAYS- 32	29.4728	5.0000	61.0000	755	0	70
BDAYS- 33	29.7728	3.0000	58.0000	744	0	81
BDAYS- 34	30.3134	3.0000	34.0000	734	0	91
BDAYS- 35	30.0650	3.0000	34.0000	723	0	102
BDAYS- 36	29.7106	8.0000	41.0000	705	0	120

BDAYS- 37	29.7161	11.0000	34.0000	641	0	184
BDAYS- 38	28.6051	8.0000	33.0000	433	0	392
BDAYS- 39	28.1304	10.0000	32.0000	46	0	779
BDAYS- 40	25.0000	14.0000	29.0000	6	0	819
BDAYS- 41	2.0000	2.0000	2.0000	1	0	824
BDAYS- 42	0.0000	0.0000	0.0000	0	0	825
BDAYS- 43	0.0000	0.0000	0.0000	0	0	825
BDAYS- 44	0.0000	0.0000	0.0000	0	0	825
BDAYS- 45	0.0000	0.0000	0.0000	0	0	825
JULMON9- 1	9.7539	7.0000	37.0000	825	0	0
JULMON9- 2	10.7745	8.0000	38.0000	825	0	0
JULMON9- 3	11.8327	9.0000	39.0000	825	0	0
JULMON9- 4	12.9370	10.0000	40.0000	825	0	0
JULMON9- 5	13.9442	12.0000	41.0000	825	0	0
JULMON9- 6	15.0461	13.0000	42.0000	825	0	0
JULMON9- 7	15.9794	14.0000	43.0000	825	0	0
JULMON9- 8	17.0352	15.0000	44.0000	825	0	0
JULMON9- 9	18.0036	16.0000	45.0000	825	0	0
JULMON9- 10	18.9964	17.0000	46.0000	825	0	0
JULMON9- 11	19.9539	18.0000	42.0000	825	0	0
JULMON9- 12	20.9758	19.0000	43.0000	825	0	0
JULMON9- 13	21.9199	20.0000	44.0000	824	0	1
JULMON9- 14	22.9915	20.0000	45.0000	824	0	1
JULMON9- 15	23.8710	21.0000	46.0000	822	0	3
JULMON9- 16	24.8902	22.0000	46.0000	820	0	5
JULMON9- 17	25.8877	9.0000	43.0000	819	0	6
JULMON9- 18	26.9451	10.0000	44.0000	819	0	6
JULMON9- 19	27.9499	11.0000	45.0000	818	0	7
JULMON9- 20	28.9242	12.0000	46.0000	818	0	7
JULMON9- 21	29.9902	13.0000	45.0000	813	0	12
JULMON9- 22	30.9815	14.0000	46.0000	812	0	13
JULMON9- 23	31.8769	15.0000	46.0000	804	0	21
JULMON9- 24	32.8294	16.0000	46.0000	803	0	22
JULMON9- 25	33.7175	9.0000	46.0000	800	0	25
JULMON9- 26	34.6884	10.0000	46.0000	796	0	29
JULMON9- 27	35.5549	11.0000	46.0000	793	0	32
JULMON9- 28	36.5217	12.0000	46.0000	784	0	41
JULMON9- 29	37.4542	13.0000	46.0000	775	0	50
JULMON9- 30	38.4306	14.0000	46.0000	771	0	54
JULMON9- 31	39.4138	15.0000	46.0000	766	0	59
JULMON9- 32	40.3894	16.0000	46.0000	755	0	70
JULMON9- 33	41.4704	17.0000	46.0000	744	0	81
JULMON9- 34	42.4537	18.0000	46.0000	734	0	91
JULMON9- 35	43.3817	19.0000	46.0000	723	0	102
JULMON9- 36	44.2411	20.0000	46.0000	705	0	120
JULMON9- 37	45.0640	21.0000	46.0000	641	0	184
JULMON9- 38	45.6374	22.0000	46.0000	433	0	392
JULMON9- 39	45.2174	30.0000	46.0000	46	0	779
JULMON9- 40	44.0000	39.0000	46.0000	6	0	819
JULMON9- 41	44.0000	44.0000	44.0000	1	0	824

JULMON9- 42	0.0000	0.0000	0.0000	0	0	825
JULMON9- 43	0.0000	0.0000	0.0000	0	0	825
JULMON9- 44	0.0000	0.0000	0.0000	0	0	825
JULMON9- 45	0.0000	0.0000	0.0000	0	0	825
HDD65N- 1	0.1550	0.0000	5.3435	825	0	0
HDD65N- 2	0.2669	0.0000	5.9082	825	0	0
HDD65N- 3	1.3625	0.0000	6.1469	825	0	0
HDD65N- 4	2.7230	0.0000	8.9099	825	0	0
HDD65N- 5	2.7498	0.0000	8.4322	825	0	0
HDD65N- 6	1.8904	0.0000	7.8291	825	0	0
HDD65N- 7	1.4154	0.0000	6.2400	825	0	0
HDD65N- 8	1.3326	0.0000	6.8299	825	0	0
HDD65N- 9	1.0554	0.0000	6.1233	825	0	0
HDD65N- 10	0.6090	0.0000	5.0026	825	0	0
HDD65N- 11	0.2128	0.0000	5.2320	825	0	0
HDD65N- 12	0.1334	0.0000	5.0026	825	0	0
HDD65N- 13	0.1453	0.0000	4.9938	824	0	1
HDD65N- 14	0.1794	0.0000	4.8313	824	0	1
HDD65N- 15	0.4267	0.0000	4.9938	822	0	3
HDD65N- 16	1.1799	0.0000	5.0026	820	0	5
HDD65N- 17	2.1674	0.0000	7.7224	819	0	6
HDD65N- 18	2.0057	0.0000	8.3562	819	0	6
HDD65N- 19	1.8618	0.0000	7.3544	818	0	7
HDD65N- 20	1.1045	0.0000	5.2421	818	0	7
HDD65N- 21	0.3356	0.0000	3.9657	813	0	12
HDD65N- 22	0.2070	0.0000	5.2421	812	0	13
HDD65N- 23	0.1457	0.0000	5.1002	804	0	21
HDD65N- 24	0.1308	0.0000	5.2421	803	0	22
HDD65N- 25	0.1552	0.0000	5.1889	800	0	25
HDD65N- 26	0.3678	0.0000	4.9187	796	0	29
HDD65N- 27	1.1705	0.0000	6.1676	793	0	32
HDD65N- 28	1.9479	0.0000	6.9948	784	0	41
HDD65N- 29	2.3993	0.0000	8.5506	775	0	50
HDD65N- 30	2.1341	0.0000	8.2429	771	0	54
HDD65N- 31	1.4714	0.0000	6.4884	766	0	59
HDD65N- 32	0.8697	0.0000	6.0484	755	0	70
HDD65N- 33	0.2707	0.0000	5.0752	744	0	81
HDD65N- 34	0.0379	0.0000	3.8860	734	0	91
HDD65N- 35	0.0236	0.0000	2.5437	723	0	102
HDD65N- 36	0.0249	0.0000	3.1488	705	0	120
HDD65N- 37	0.0284	0.0000	2.7164	641	0	184
HDD65N- 38	0.0889	0.0000	3.3928	433	0	392
HDD65N- 39	0.1195	0.0000	1.8627	46	0	779
HDD65N- 40	0.1383	0.0000	0.6652	6	0	819
HDD65N- 41	0.0000	0.0000	0.0000	1	0	824
HDD65N- 42	0.0000	0.0000	0.0000	0	0	825
HDD65N- 43	0.0000	0.0000	0.0000	0	0	825
HDD65N- 44	0.0000	0.0000	0.0000	0	0	825
HDD65N- 45	0.0000	0.0000	0.0000	0	0	825
survey variables:						

SQFT	2770.0963	100.0000	9999.0000	405	0	420
YRBUILT	68.2954	49.0000	96.0000	369	0	456
REBEQUIP	0.9514	0.0000	1.0000	329	0	496
DREPLAPROG	0.3321	0.0000	1.0000	825	0	0
DADDITPROG	0.0618	0.0000	1.0000	825	0	0
OLDPGGAS	0.3188	0.0000	1.0000	825	0	0
OLDPGELE	0.0085	0.0000	1.0000	825	0	0
SURPRISE	0.0109	0.0000	1.0000	825	0	0
NEWCOOK	0.1270	0.0000	1.0000	795	0	30
DREPLA	0.0921	0.0000	1.0000	825	0	0
DADDIT	0.0400	0.0000	1.0000	825	0	0
OLDGAS	0.0764	0.0000	1.0000	825	0	0
OLDELE	0.0121	0.0000	1.0000	825	0	0
GASINCR96	0.0668	0.0000	1.0000	793	0	32
GASDECR96	0.0151	0.0000	1.0000	793	0	32
GASPCTCHG	23.3684	1.0000	100.0000	57	0	768
SRVBREAK	1.0000	1.0000	1.0000	318	0	507
SRVLUNCH	1.0000	1.0000	1.0000	636	0	189
SRVDINNER	1.0000	1.0000	1.0000	610	0	215
SRVNOMEALS	1.0000	1.0000	1.0000	28	0	797
SRVOTHER	1.0000	1.0000	1.0000	22	0	803
PCTBREAK	25.3904	0.0000	100.0000	292	0	533
PCTLUNCH	47.4610	10.0000	100.0000	603	0	222
PCTDINNER	44.8243	0.0000	100.0000	649	0	176
NUMMEALS	2244.6521	0.0000	10000.0000	434	0	391
MEALSMORE	0.5645	0.0000	1.0000	248	0	577
MEALSLESS	0.4355	0.0000	1.0000	248	0	577
NMEALS95	2129.0440	0.0000	10000.0000	91	0	734
MTFAST	1.0000	1.0000	1.0000	383	0	442
MTPIZZA	1.0000	1.0000	1.0000	83	0	742
MTSELF	1.0000	1.0000	1.0000	35	0	790
MTTABLE	1.0000	1.0000	1.0000	322	0	503
MTOTHER	1.0000	1.0000	1.0000	102	0	723
DAYSPERWK	6.7632	1.0000	7.0000	815	0	10
YEARROUND	0.9804	0.0000	1.0000	816	0	9
CLOSE01	0.2500	0.0000	1.0000	16	0	809
CLOSE02	0.2000	0.0000	1.0000	15	0	810
CLOSE03	0.2000	0.0000	1.0000	15	0	810
CLOSE04	0.2000	0.0000	1.0000	15	0	810
CLOSE05	0.2000	0.0000	1.0000	15	0	810
CLOSE06	0.3333	0.0000	1.0000	15	0	810
CLOSE07	0.4000	0.0000	1.0000	15	0	810
CLOSE08	0.4667	0.0000	1.0000	15	0	810
CLOSE09	0.2667	0.0000	1.0000	15	0	810
CLOSE10	0.2000	0.0000	1.0000	15	0	810
CLOSE11	0.2000	0.0000	1.0000	15	0	810
CLOSE12	0.4000	0.0000	1.0000	15	0	810
SAMELEVEL	0.6241	0.0000	1.0000	798	0	27
HIVOL01	0.1752	0.0000	3.0000	314	0	511
HIVOL02	0.1465	0.0000	1.0000	314	0	511

HIVOL03	0.1178	0.0000	1.0000	314	0	511
HIVOL04	0.0987	0.0000	1.0000	314	0	511
HIVOL05	0.1210	0.0000	1.0000	314	0	511
HIVOL06	0.3312	0.0000	1.0000	314	0	511
HIVOL07	0.3503	0.0000	1.0000	314	0	511
HIVOL08	0.3503	0.0000	1.0000	314	0	511
HIVOL09	0.1592	0.0000	1.0000	314	0	511
HIVOL10	0.1115	0.0000	1.0000	314	0	511
HIVOL11	0.3726	0.0000	1.0000	314	0	511
HIVOL12	0.4936	0.0000	1.0000	314	0	511
BCMORE	0.9655	0.0000	1.0000	29	0	796
BCLESS	0.0000	0.0000	0.0000	0	0	825
BCRENOV	1.0000	1.0000	1.0000	43	0	782
BCSAME	1.0000	1.0000	1.0000	707	0	118
BCMORESF	755.5862	1.0000	3500.0000	29	0	796
BCLESSSF	0.0000	0.0000	0.0000	0	0	825
COOLFUEL	3.2368	1.0000	9.0000	815	0	10
HEATPCT	77.4906	0.0000	100.0000	795	0	30
HEATGAS	0.7091	0.0000	1.0000	629	0	196
HEATNEW95	0.0750	0.0000	1.0000	693	0	132
HEATREPLA	0.7647	0.0000	1.0000	51	0	774
HEATWASGAS	0.5938	0.0000	1.0000	32	0	793
HEATNEWYR	2.1600	1.0000	3.0000	50	0	775
HEATNEWMO	6.3415	1.0000	12.0000	41	0	784
HEATNEWPCT	76.9362	5.0000	100.0000	47	0	778
WHEATGAS	0.9152	0.0000	1.0000	743	0	82
WHEATNUM	1.2349	0.0000	10.0000	779	0	46
WHEATCAP	107.3212	20.0000	1000.0000	495	0	330
WHEATNEW95	0.1335	0.0000	1.0000	764	0	61
WHEATWASGA	0.9192	0.0000	1.0000	99	0	726
WHEATNEWYR	2.2121	1.0000	3.0000	99	0	726
WHEATNEWMO	6.4444	1.0000	12.0000	72	0	753
WHEATREPLA	0.9406	0.0000	1.0000	101	0	724
WHEATPCT	53.5714	10.0000	90.0000	14	0	811
DIWASHCHG	0.1281	0.0000	1.0000	773	0	52
DIWASHMORE	0.7576	0.0000	1.0000	99	0	726
DIWASHLESS	0.2424	0.0000	1.0000	99	0	726
OTHERGAS	0.0101	0.0000	1.0000	792	0	33
SAMPCELL	9.7205	1.0000	13.0000	347	0	478
ANALYWGT	3.7061	2.1667	10.7222	347	0	478

M&E PROTOCOLS TABLE 6 - RESULTS FOR STUDY ID 711

DSM Program = Commercial Equipment Replacement, End-Use Element = Cooking

(Designated Unit of Measurement: Load Impact Per Installation[a])

				90 % Confidence				80% Confidence			
				Lower Bnd	Upper Bnd	Lower Bnd	Upper Bnd	Lower Bnd	Upper Bnd	Lower Bnd	Upper Bnd
1. Avg Participation Group and Avg. Comparison Group		**Part Grp**	**Comp Grp**	**Part Grp**	**Part Grp**	**Comp Grp**	**Comp Grp**	**Part Grp**	**Part Grp**	**Comp Grp**	**Comp Grp**
A.Base usage:	Pre-installation therms	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
	Base therms	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
	Base therms/designated unit	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
B.Impact year usage	Impact yr therms	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
	Impact yr therms/designated unit	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
2. Avg. Net and Gross End Use Load Impacts		**Avg Gross**	**Avg Net**	**Avg. Gross**	**Avg Gross**	**Avg Net**	**Avg Net**	**Avg Gross**	**Avg Gross**	**Avg Net**	**Avg Net**
A.Load Impact[b]	Load impact: annual therms x 1,000	-138	546	---[d]	---[d]	464	694	---[d]	---[d]	486	672
B.Load Impact/unit	Load impact/Installation: annual therms	-201[c]	793	---[d]	---[d]	674	1,007	---[d]	---[d]	706	975
C.Pct Chg in usage	Pct chg - Part grp - therms	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
	Pct chg - Comp grp - therms	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*	*opt.*
D. Realization rates	Load impact: therms - realization rate	-0.20[e]	1.06	---[d]	---[d]	0.85	1.27	---[d]	---[d]	0.89	1.23
3. Net-to-Gross Ratios		**Ratio**	-	**Ratio**	**Ratio**	-	-	**Ratio**	**Ratio**	-	-
A.Avg impact	Average load impact - therms	-3.95		n/a	n/a			n/a	n/a		
B.Avg impact/unit	Avg load impact/Installation - therms	-3.95		n/a	n/a			n/a	n/a		
C.Avg load impact,	Pct chg, impact yr. rel. to base yr: therms	*opt.*		*opt.*	*opt.*			*opt.*	*opt.*		
4. Designated Unit Intermediate Data		**Part Grp**	**Comp Grp**	**Part Grp**	**Part Grp**	**Comp Grp**	**Comp Grp**	**Part Grp**	**Part Grp**	**Comp Grp**	**Comp Grp**
A. Pre-installation	Pre-installation average value	---	---	---	---	---	---	---	---	---	---
B. Post-installation	Post-installation average value	---	---	---	---	---	---	---	---	---	---

6. Measure Count Data	**Number**
A. # Measures, installed by participants in "Participant Group."	347
B. # Measures, all pgm participants, 12 mo of pgm yr.	689
C. # Measures, installed by "Comparison Group"	---
7. Market Segment Data	**Number**
A. Number of Participants	689

[a] The designated unit of measurement for this program is the overall per-customer installation, including all rebated equipment items together.

[b] All total annual impacts in this row were calculated by multiplying the impact/installation by the total number of participants.

[c] The gross impact is negative because the cooking component of the model increased in 1997, the post-program period.

[d] The confidence intervals for the gross impact were not computed. Please refer to the relevant discussion following equation IV.10 in Section IV of the report.

[e] The negative gross impact results in a negative realization rate and net-to-gross ratio also.

Table 7 – Implementing The Documentation Protocols For Data Quality And Processing

This document is organized to follow the exact structure of Table 7 of the Protocols. The section headings and bold text found throughout have been copied from that source so that the information provided by the research team can be seen in context.

A. Overview Information

A.1. Study Title and Study ID No.: The study title and identification number should be identical to the information contained in the Statewide Bibliography. Changes in this information should be noted.

First Year Load Impacts of Southern California Gas Company's
Program Year 1996
Commercial Energy Efficiency Incentive Program

Study # 711

A.2. Program year (or years) and program description: The program and program year(s) should be identical to the information contained in the Statewide Bibliography.

The primary emphasis of the 1996 SoCalGas ER DSM program was upon natural gas savings for cooking. Under this program SoCalGas customers were offered rebates for installing relatively efficient natural-gas-using equipment, primarily, though not exclusively for cooking. The targeted market segments for this program were SoCalGas commercial customers who were classified as belonging to the following building (business) types.

1. Large office
2. Small office
3. Restaurant
4. Large retail
5. Small retail
6. Food stores
7. Refrigerated warehouses
8. Unrefrigerated warehouses
9. Hotels and motels
10. Primary and secondary education
11. Colleges and universities
12. Hospitals and clinics
13. Miscellaneous commercial

The program participants, however, tended to be concentrated in the restaurant and education and food store building classifications. The program was first offered in January 1996 and extended to December 1996.

A.3. End-Uses and/or measures covered: Use the end use designations agreed to in the Protocols.

The SoCalGas equipment replacement effort was primarily directed at cooking although some insulation, HVAC, and water-heating measures were also included in the program. The non-cooking rebates were relatively few in number (a total of 11) and in this evaluation will be included in the miscellaneous measures.

A.4. Methods and models used: Describe the final model specification used for the study. Where applicable, indicate the report location of the competing class or types of models that were estimated but were not selected.

Although we estimated a series of preliminary sets of regressions, the early models were designed for simple experimentation with a complex new data set rather than as specifications that would account properly for the effects we wished to measure. We did have several ideas for alternative approaches but regulatory time limits precluded investigation of them. In this section of the table we will summarize the equation specification presented in section III of the paper. A further set of specification issues will be discussed below in connection with the issues listed as a., b., c. and d. under the model specification heading in Table 7 of the Protocols.

The equation specification upon which this regression is based is given in Section III of the study text.

A.5. Participant and comparison group definition: Specifically present the study's definition of what constitutes a participant and comparison group member. Refer to Table 5 of the Protocols, where appropriate.

Participants for this study consist of SoCalGas Customers who purchased at least one item of cooking equipment that qualified as a rebate measure during the course of the 1996 SoCalGas Commercial Equipment Replacement DSM program year. To be included as an official participant, the program tracking records had to show evidence of a greater that zero dollar payment made for at least one measure associated with a completed customer application.

Non-participant Comparison Group was defined as active SoCalGas Commercial customers who had some type of cooking facility at their establishment and did not participate (receive a rebate) as part of the 1996 SoCalGas Commercial Equipment Replacement DSM program. Customers were chosen randomly from a list of businesses in the high frequency building types observed for program participants. DSRA analysts matched program participation records to these so that the set of non-matches could be used as the non-participant frame.

A.6. Analysis sample size: Provide the number of customers, number of installations, number of measures (if different) and the number of observations (monthly units of analysis) if different for different units of analysis, a summary table should be provided.

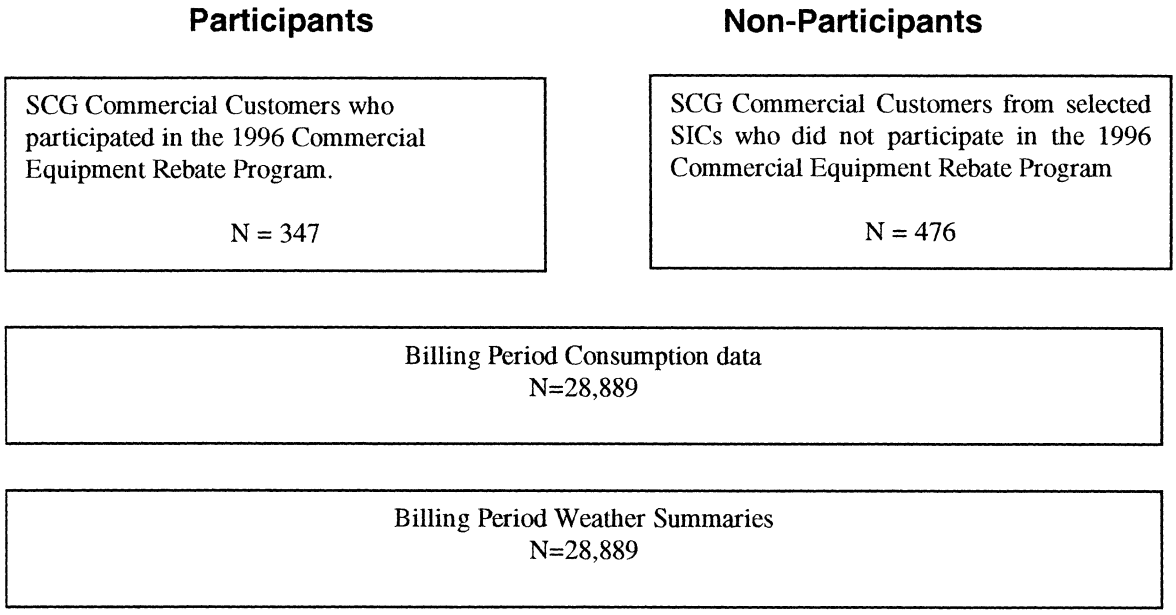
The number of customers in the analysis sample size was 823 consisting of 347 participants and 476 nonparticipants. The total number of monthly billing observations on these customers was 28,889.

B. Database Management

B.1. Describe and provide a flow chart illustrating the relationships between data elements.

Figure B.1. shows the sources of information that were used for both participant and non-participant customer groups used in the final analytic dataset for this study. Figure B.1. should be viewed from the top down within each of the columns, labeled participants or non-participants.

Figure B.1.
Participant and Non-Participant Evaluation Data Structure



The bottom of Figure B.1. shows that for all customers used for this evaluation, monthly billing period consumption data was assembled and matched to billing period weather summaries. In total, 28,889 monthly billing period consumption and weather observations were matched to participant and non-participant customers.

B.2. Data sources

Table B.2. provides additional detail about the source of data items used for the commercial evaluation project. The left column identifies types of data and particular variables that were used for different parts of the analysis evaluation project. The right column identifies the specific data set that was the source of these items. Figure B.1. and

Figure B.2.
Sources of Evaluation Data

Data Item	Source
Billing Data: from and to billing period dates, total billing days and billing-period Therms.	SCG's Customer Consumption Billing file provided for selected customers by SoCalGas and MCGI
Weather Data: weather zone ID Value, heating degree days, cooling degree days, average temperature.	SCG's Weather File pre-merged to individual billing period records by SoCalGas and MCGI analysts.
Program Participant Data: location numbers, and rebate indicators.	SCG's 1996 Commercial Equipment Rebate Program participation tracking databases.
Survey Response Data: Individual survey response values for all program participant and non-participant responses.	Output from DSRA's Computer Assisted Telephone Interviewing system. These data were transferred to SAS datasets.
Weights: survey respondent response weights constructed for business size and type strata.	Decision Sciences' response-weighting dataset.

Program Participant Data The program tracking system is organized by application and measure. It contains information about the characteristics of the rebated measure and the customer. Important variables include a summary description of the measure, the amount and date of rebate payment.

Billing and Weather Data. Monthly billing data for these samples covered the period from January 1995 through October 1997. These data were provided for all customers in the final analysis dataset. Summary weather variables (heating degree days and cooling degree days and average temperature) were appended to the individual billing data by SoCalGas analysts, contemporaneous with the billing periods; these were based on readings at the weather stations closest to the business location.

B.3. Diagram and describe the data attrition process commencing with the program data-base for participants, and the utility database for the comparison group. Specific numbers and decision points for inclusion and exclusion should be provided. Where different data sources are used (e.g., surveys and program records), appropriate attribution categories should be used (e.g., response rates for surveys, unidentified account numbers for participants' billing records).

Figure B3 – Distribution and Selection of Program Participation Records

Customer Applications	
Total (*unique account numbers*)	1,505
Match at least one measure record	1,290
Do not match any measure records	215
Measure Records	
Total (*all match customer application records*)	1,784
Cooking Measures	1,375
Valid - payment > $0.00	965
Incomplete	410
Non-Cooking Measures	409
Valid - payment > $0.00	238
Incomplete	339
Program Participants	
Total (with 1 or more Cooking Measures)	689
Sum of all Measures	976
Sum of Cooking Measures	965
Sum of Non-Cooking Measures	11

B.4. Describe the internal/organizational data quality checks and data quality procedures used to match customers and billing records, surveys, participation records, weather data, and any other data used in the analysis.

Data used in this analysis was identified using the SoCalGas variable values assigned to track customers and program participants. Gas Company customers are identified by both Account and Premise ID numbers. These values were found on program participation records providing the match between historical billing records and program participation status variables.

Once data were delivered to the analysis team, DSRA assigned a unique Case Identification Number (CASEID) that served as the unchanging referent to the SoCalGas

customer. The case IDs were used to provide the sample candidate identification for DSRA's Computer Aided Telephone Survey system and to match to samples and files of consumption values. The Case ID values followed the survey response dataset to AEI for purposes of matching response data to both weather and billing data.

Note that during the study year, 1997, various SoCalGas customer identification values were administratively changed. This made obsolete, some of the references found in the program tracking records because they referred to pre-program customer ID values. Subsequent consumption data was supplied by SoCalGas and MCGI analysts in November, 1997 using a correspondence files and rules to match old to new identifiers.

B.5. Provide a summary of the data collected specifically for the analysis but not used, the reasons for them not being used, and a documentation of where those data reside.

For this analysis we used only program participation records that were associated with valid program participants. Other records of incomplete applications or measures not qualified for rebates were excluded. In addition, our telephone survey was designed to provide a general characterization of the candidate businesses and their commercial cooking activities. Not all of these survey items were used for the construction of the savings mode because we chose a more aggregate EUI model to assess the program impacts We collected these additional items, to insure a logical survey script and to provide data for additional analysis of businesses with commercial cooking activities.

The Survey Instrument is appended to this document. A review of the instrument will show that it contains very detailed information about the SoCalGas Company customers that will be very useful for company planning. Our approach to using this information in our modeling effort was to balance the complexity of the model against the desire for explanatory power. Our goals were as follows: we tried to keep the model as simple as we could and still account for the major theoretical issues that faced us.

.

C. Sampling

C.1. Sampling procedures and protocols: Describe the sampling procedures and protocols used. Information provided should include the sampling frame (i.e., eligible population), sampling strategy (e.g., random, stratified, etc.), sampling basis (e.g. customers, installation, rebate issued), and stratification criteria (e.g., geographic, etc.). Descriptions should be provided separately for participant and comparison group data. Specific data and formulas should be used to present sampling goals and achieved results. Procedures to calculate sample sizes in order to achieve specific levels of precision at given levels of confidence should be explicitly described.

SCG Commercial Equipment Replacement Participant Sample: The telephone survey was directed at customers who had participated in the 1996 SCG Commercial Equipment Rebate Program. The gross sample for this survey was drawn by choosing all program participant records for which phone numbers could be obtained. This consisted of a total of 689 unique program participant locations Technically, the sample for the participant of study was an attempt census all available program participants.

Non-participant Survey: 2,000 sample candidates were randomly selected from a SoCalGas file of all commercial businesses defined by the high frequency building types observed for program participants. This file was first matched to program participation records using both computer matching and visual inspections to insure that valid program participants would not be included. No other pre-survey strata were defined for non-participants chosen for use as candidates for the comparison group. Our plans provided for the fact that after the survey was conducted we would obtain late 1997 consumption data for these candidates for possible post hoc stratification if significant differences in business size could not be accounted for using the estimation model itself.

Analysis Sample Weighting

Response sample weights were constructed for respondent program participants. These were based on post-hoc stratification of participants into three size categories for two types of businesses; restaurants and all others. This strategy produced response weights designed to correct for the differential response by business size and type. These weights and similar ones calculated for non-respondents are not employed in the evaluation presented here because the estimation model itself accounts for differences in business type and size.

C.2. Survey information: Survey instruments should be provided Response rates should be presented. Reasons for refusals should be presented in tabular form. Efforts to account for or test for non-response bias should be presented, as well as corrections to account for the bias.

The table below presents a summary of the response experience for the *1996 SCG Commercial Equipment Replacement Evaluation Survey.* Observe that of the 2,689 sample members, a total of 2,296 contacts were attempted and a total of 850 valid and complete responses were obtained.

Figure C2 – Non-Participant Follow-up Survey Sample Accounting Summary

	Count	Percent
Sample	2,689	
Total Attempted Contacts	2,296	85.4%
Total Phone Calls Placed	5,752	
Non-contact		
Not Called	393	14.6%
Not In Service	460	20.0%
Busy	17	0.7%
No Answer	6	0.3%
Answering Machine / Voice Mail	38	1.7%
Wrong Number	17	0.7%
Other phone problems	3	0.1%
Contact		
Requested callback	535	23.3%
Refused	223	9.7%
Unable to provide information	42	1.8%
No knowledgeable respondent	105	4.6%
Complete Interviews	850	37.0%

C.3. Statistical descriptions: For the key variables that were used in the final models, provide descriptive statistics for the participant group, and, when present for the comparison group.

In our response to this item, we present the unweighted means of the variables in the data set that generated the final models. The details of the screening filters applied to this data set are given in our response to item D.3. below.

D. Data Screening And Analysis

D.1 Describe the procedures for the treatment of outliers, missing data points and weather adjustment.

Outliers: There were no observations that were excluded using a poor fit rationale. Although the exclusion of outliers can be useful, it must also be used with discretion since we would not want to eliminate valuable sample information, unnecessarily. In the current case, we could not assure ourselves that we would not be throwing away valuable information if we eliminated outliers.

We also eliminated the first billing period observation for each customer since we were using lagged values of our variables in our model. In addition we eliminated observations if the therm level was less than 10 therms per month since these levels most likely indicated a simple pilot light load, rather than an energy consumption process that included actual cooking.

Missing Data Points: The data set was organized as a cross-section/time-series file. In this data set, there were up to three years of monthly data for each of the customers in the sample. Observations were eliminated at the customer level if we lacked the following information about the customer:

- business (building) type
- program participation status
- a valid SoCalGas customer location

Observations were eliminated at the monthly level if we lacked information on billing or weather data for the month. Note that observations eliminated at the customer level would result in elimination of all the monthly observations for a customer.

Weather Adjustment: The weather variables used in the model were cooling and heating-degree-days. Normalized values for these variables were constructed by calculating a three-year average weather pattern for the weather station corresponding to each customer's billing record. Although the recorded weather data was used in the regression equation, the recorded weather values were set equal to the corresponding normalized values for the purposes of calculating the weather-normalized impacts and consumption levels.

D.2. Describe what was done to control for the effects of background variables, such as economic and political activity that may account for any increase or decrease in consumption in addition to the DSM program itself.

We used a difference-of-differences (DOD) method as outlined in Table 5 of the Protocols. The primary argument for using this method is that it has been thought to account for the effects of background variables common to participants and nonparticipants by giving credit for program savings *only* for the excess of participant impacts over nonparticipant impacts. In addition, we accounted for any general changes in consumption for participants and nonparticipants for each of our three sample years (1995-1997). Of course, given enough time it would be of intellectual interest, and probably of use for forecasting and planning, to include variables summarizing economic and political activity in the regression model. Given our use of the DOD method, however, it is unclear whether this would affect our savings estimates noticeably.

D.3. Describe procedures, including those identified in Table C-12 used to screen data for inclusion into the final analysis data set. Show how many customers, installations or observations were eliminated with each screen. The reviewer should be able to clearly follow the development of the final analysis dataset.

In section B.3 we described the data attrition process up to the construction of the *final analytical data set.* In the following table we present the data screening process applied to this final data set. The analysis was performed on a pooled cross-section/time-series

data set and there were time-series and cross-section screening criteria applied to the data. We will consider these separately, to the extent possible.

Time-series observations were dropped only for missing billing data. These observations were dropped because of the importance of the amount and timing of consumption in our analysis. Of course, the *amount* of consumption is important since it is the dependent variable in our regression equation. The *timing* of consumption is just as important an issue since the Protocols impact calculations require that we have a reasonably clear assignment of each billing period to the pre-program-year, program-year, or post-program-year periods.

Customer-level observations were dropped from the analytical data set lack of a location identifier. The location identifier was important because it assured us that a correct match with the billing readings of consumption could be obtained. In addition, the business/building-type of the customer (and, therefore, the relevant set of EUI's) was tied to this variable. The program participation status was also used as a screening variable in our original work, but no observations were screened because of this variable.

D.4. Regression statistics: For all final models, provide standard regression statistics in a tabular form.

A table of regression results is present as Table 3.1 in Section III of the study.

D.5 Specification: Refer to the section(s) of the Study that present the initial and final model specifications that were used, the rationale for each, and the documentation for the major alternative models used.

The discussion of alternative approaches is contained in Appendix I, Section A.I.3.

D.5.a. describe how the model specification and estimation procedures recognize and address heterogeneity of customers (i.e., cross-sectional variation)

The characteristics of the customers differ in our sample and we take account of variation associated with differing business (building) types, variations in floor space and differing participation status.

D.5.b. describe how the model specification and estimation procedures recognize and address changes in factors that affect consumption over time (i.e., time series variation), apart from program effects

The model accounts for time-related changes in basic consumption for each of the three years in our sample (1995-1997). Furthermore the model accounts for time-related variation in the recorded weather as described in Appendix I of the study. The following

summary of the treatment of the weather variables follows the discussion in that Appendix.

We also eliminated the first billing period observation for each customer since we were using lagged values of our variables in our model. In addition we eliminated observations if the therm level was less than 10 therms per month since these levels most likely indicated a simple pilot light load, rather than an energy consumption process that included actual cooking.

Since we expect the loads for space heating to be weather-sensitive, we create interaction terms to adjust the heating EUI's for the effects of changing weather conditions. This weather variable is constructed as a ratio to normal weather so that they affect the EUI's only to the extent that current weather conditions differ from normal weather. In the remaining discussion then we will use the following equation when referring to the space heating EUI.

A.I.1.4) $$EUI_{h,bt} = (ncdd_m \cdot EUI_{ac,bt} + nhdd_m \cdot EUI_{sh,bt})$$

Table A.I.3
Definitions Of New Terms In Equation A.1.4

Term	Units	Definition
$EUI_{ac,bt}$	monthly Thm /sqft	energy use index for air-conditioning in building-type *bt*
$ncdd_m$	pure number	*recorded* monthly cooling degree-days divided by *normalized* monthly cooling degree-days, in month *m*
$EUI_{sh,b}$	monthly Thm /sqft	energy use index for space-heating in building-type *bt*
$nhdd_m$	pure number	*recorded* monthly heating degree-days divided by *normalized* monthly heating degree-days, in month *m*

D.5.c. describe how the model specification and estimation procedures recognize and address the fact that participants self-select into that status, and discuss the effects of self-selection on model estimates whether or not self-selection is treated explicitly.

A selectivity correction term could be used to address the self-selection bias issue and this term would be modeled as a function of cross-customer, not time-series, characteristics. This means that such a term would cancel and thus be unecessary in our first-difference regression. Since the first-difference regression was therefore corrected for any self-selection bias and since the first-difference form resulted in a larger estimate of the program impact coefficient, there seemed to be no compelling argument to include a selectivity correction term in the regression model presented in Section III.

D.5.d. discuss the factors, and their associated measures, that are omitted from the analysis and any tests, reasoning, or special circumstances that justify their omission

We did not knowingly omit consideration of any factors from the analysis that were called for in the Protocols. There are for any study, however, an infinite number of factors that *might* be important. Our approach to this problem is to account for the effects of the factors that seem of most importance first, and to continue testing any other promising factors as the regulatory schedule permits.

D.5.e. describe how the model specification can be interpreted to yield the net impacts measure(s) recognized in Section E below.

We addressed the issues of the relationship between the model specification and the net impact measurements in Section IV.A of the study.

.

D.6. Error in measuring variables: Describe whether and how this issue was addressed and what was done to minimize the problem.

Throughout an evaluation project of this complexity, measurement error can occur at many points. Previously, in section B.4., the procedures used to insure valid matches using SoCalGas program files were described. These procedures served to include in the analysis only those survey consumption and weather data records for which a certain match key was available. Survey are, of course, subject to various types of error, most notably respondent error, that may arise from self-reported respondent misunderstanding of a question or providing an answer for which they are not qualified. Since the surveys used for this analysis all consisted of telephone interviews, the normal respondent error that might be expected from mailout/mail back instrumentation should be absent.

Telephone interviewers were trained and supervised so that they would prompt respondents if answers were unclear or seemed inappropriate for the questions asked. In the case where answers were suspect during the telephone interviewing, interviewers asked if another knowledgeable person was available who could help with areas of uncertainty. Interviewers were also trained to enter "missing" or "did not know" when problems could not be resolved.

We are confident that very low levels of error occurred with regard to the assignment of billing and weather data to the subject customer locations. Previously, in section B.4., the procedures for matching weather data to customer location were described. No cases were included in the final analysis set for which weather data could not be attributed, and none of the weather station data covered during the study period included missing observations.

Billing data is subject to some level of problems associated with miscalculated or estimated bills. SoCalGas's files contain billing codes that indicate whether or not a

particular billing consumption period was estimated or an adjustment for a previous observation dataset, so that billing idiosyncrasies would not add erroneous variation to the monthly values. These codes were reviewed and implemented by SoCalGas and MCGI analysts prior to delivering the billing data to the research team.

D.7. Autocorrelation problems and the solutions specifically taken to address the problem. Specific identification and mitigation diagnostics should be presented including differing treatment for sub-groups, if any.

We used a generalized difference estimation routine that made use of an estimated serial correlation coefficient. The procedure is described in Appendix I, Section A.I.b of the study report.

D.8 Heteroskedasticity: Describe the diagnostics carried out, the solutions attempted and the effects. If left untreated, explain why.

The heteroskedasticity issue was addressed by estimating the impact regression equation using weighted-least-squares. The procedure is described in Appendix I, Section A.I.b of the study report.

D.9 Collinearity: Describe procedures used to address the problem of collinearity, and the reasons for either not treating it or treating it to the level that was done.

The collinearity issue was not a major concern in our study.

D.10 Influential Data Points: Describe the influential data diagnostics that were used, and how the identified outliers were treated.

This issue is related to the discussion in our response to Protocol issue D.1. above. There were no observations that were excluded using a poor fit rationale. In the current case, we could not assure ourselves that we would not be throwing away valuable information if we eliminated outliers.

D.11. Missing data: Describe the methods used for handling missing data during the analysis phase of the study.

This issue is related to the discussion in our response to Protocol issue D.1. above. As we noted in that response, the data set was organized as a cross-section/time-series file. In this data set, there were up to three years of monthly data for each of the customers in the sample. Observations were eliminated at the customer level if we lacked the following information about the customer:

- business (building) type
- program participation status
- valid SoCalGas customer ID number

Observations were eliminated at the monthly level if we lacked information on billing or weather data for the month. Note that observations eliminated at the customer level would result in elimination of all the monthly observations for a customer.

The following discussion of missing floorspace data is taken from Appendix I of the study report.

The building types considered in the SoCalGas forecasting model are summarized in appendix A.II.1. In our sample, however, we made a further categorization of customers depending upon the floorspace information available for them. There were three categories of floorspace information: 1) the floorspace specified by the survey respondents, when that was less than 10,000 square feet; 2) a 0,1 dummy variable that has the value 1 only if the floorspace was greater than 10,000 square feet; and, 3) a 0,1 dummy variable that has the value 1 if the floorspace information is entirely missing. To account for these three categories we estimate a different coefficient for each building-type/square-footage category. For example, we estimate different α_{bi} coefficients for restaurants with specified and missing floorspace. As noted in section III, for respondents with no specified floorspace, the estimated coefficients account for differences in floorspace as well as differences in the applicability of the prior EUIs.

In our regression model, then, the BT summation limit in equation A.I.1.5b takes into account all the original SoCalGas building types and the three categories of floorspace information. The complete list of these α_{bi} is included in table III.1, *Regression Results*, in section III. To implement this specification we replace the FS variable above with the following expression,

A.I.1.8) $FSterm = FSspec + d_{miss} + d_{10k}$

where, *FSspec* is the actual floorspace (in square feet) where it is specified, and *FSspec* is 0 otherwise; d_{miss} is a 0,1 dummy variable that is 1 if the floor space is not specified, and d_{miss} is 0 otherwise; and, d_{10k} is a 0,1 dummy variable that is 1 if the customer's floor space is greater than 10,000 square feet, and d_{10k} is 0 otherwise.

D.12. Precision: Present the methods for the calculation of standard errors for key parameters such as gross impacts, net impacts, and net-to-gross ratios.

The standard errors for the net effects are relatively straightforward. The standard errors for the gross impacts are somewhat complex, however, because of the nonlinearities in the estimating equation.

To summarize these calculations, the net impact confidence intervals were calculated using the standard approach. The gross impact confidence intervals, however, should be considered only approximate. Note, however, that all the t-ratios involved are relatively large.

E. Data Interpretation And Application

E.1. For all program participants and at the end use level, net impacts should be calculated as either:

E.1.a. Average participant group load impacts minus average comparison group load impacts, plus or minus the effects of uncontrolled differences between the participant and comparison groups times the number of participants;

This was the method we used. Our application of the method is described above under our response to item D.5.e., as well as Section IV of the Study text.

E.1.b. Average participant group load impacts per designated unit, times the number of designated units, times the NGR (based on designated units);

E.1.c. (If comparison group load impacts not available) Average participant group load impacts, times the number of participants times the NGR (derived independently), or;

E.1.d. Other application methods agreed upon by the CADMAC.

E.2. Describe the process, choices made and rationale for choices made in Section E.1 above.

The method we used was that described under item E.1.a. above. We chose this method because we had nonparticipant data available and we were able to specify the analysis so that the Protocol Difference of Differences was estimated as a term in the regression equation (again, see the discussion under item D.5.e., above, as well as section IV of the Study text. As we noted in Section III of the Study text, the regression equation is designed to facilitate a comparison of participant vs. nonparticipant consumption changes, taking into account differences in customer floor space as well as the commercial building type, changes in weather, and the presence of installed DSM measures associated with the 1996 SoCalGas equipment replacement program.

SCG Commercial Equipment Replacement Evaluation Survey

Hello. My name is _____, and I am calling on behalf of The Gas Company

Would you mind giving us a few minutes of your time to answer questions about the your business facility and the types of cooking equipment you use.

The information that you share is going to be used to help The Gas Company plan for the future and provide better service and value to its customers. Your answers will be kept strictly confidential.

First, we would like to ask you a few questions about your business activity and the characteristics of the building where your business is located.

Establishment Site Activity

>001< What type of business activity is conducted at this location?

OFFICE	RESTAURANT	FOOD STORE
11 Admin. and management	21 Fast food/self-serv.	31 Supermarket
12 Financial/Legal	22 Table service	33 Other food store
13 Insurance/Real Estate	23 Bar, nightclub, other	
14 Other office		
HEALTH CARE	EDUCATION	LODGING
61 Hospital	71 Daycare or preschool	81 Hotel
62 Nursing home	73 College, university	82 Motel
63 Medical office	74 Vocational or trade	
64 Clinic	75 K-12 Schools	
PUBLIC ASSEMBLY	OTHER	
91 Church	0 Specify	
92 Recreational/other	90 Retail	

Answer ===>

>002< Do you own or rent the space your business occupies?

<1> Own/buying
<3> Lease/rent
<5> Don't know/Won't say

Answer ===>

>003< How long has your business/ organization been at the current location?

<1-50> years
<99> don't know

Answer ===>

>004< Does this business occupy the entire building or is the building shared with other businesses?

<1> Occupies entire building
<3> Shares building with other businesses

Answer ===>

>005< What is the number of floors in this building?

<1-20> Floors
<21> High rise, with more than 20 floors
<99> Don't know

Answer ===>

>006< Is your establishment on the ground floor, top floor or between the ground floor and top floor?

<1> Ground floor
<2> Top floor
<3> Between the ground floor and top floor
<5> Don't know/Won't say

Answer ===>

>007< How many square feet does your facility occupy?

<100-99999> total square feet
(use 99999 for 100,000 sq ft or more
<99> Don't know, Won't say

Answer ===>

>008< Do you know the year the building your business occupies was built?

<49>> before 1950
<50-96> 1950 to 1995
<9> Don't know, Won't say

Answer ===>

>participant<

>009p1< We understand that some Gas Company Rebate Program equipment is installed at this location, is that correct?

<1> Yes
<3> No [goto thnk]
<5> Don't Know [goto thnk]

Answer ===>

>009p2< Was the new equipment purchased to replace old equipment or was it purchased as additional equipment?

<1> Purchased as replacement equipment
<2> Purchased as additional equipment only[goto 010p]
<3> Both replacement and additional
<5> Don't know [goto 010p]

Answer ===>

>09p3< What type of fuel did the replaced equipment use?

<1> Gas
<3> Electric
<5> Don't know

Answer ===> [goto 010p]

>nonparticipant< continue

>09n1< Are you aware that the Gas Company offered a rebate program for commercial gas cooking equipment in 1996?

<1> Yes
<3> No [goto 010n]
<5> Don't know [goto 010n]

Answer ===>

>09n2< Did you take part in the Gas Company's 1996 Equipment Replacement Program? (and receive a rebate)

<1> Yes [goto thank you]
<3> No
<5> Don't know

Answer ===>

Now we would like to ask you a few questions about the types of cooking equipment you use at your business.

>010n< Have you installed any new or used cooking equipment since January 1, 1996?

<1> Yes
<3> No [goto 013]
<5> Don't know

Answer ===> [goto 011]

>010p< [equiv 010n] *Now we would like to ask you a few questions about the types of cooking equipment you use at your business.*

Have you installed any new or used cooking equipment since January 1, 1996? (other than Program Rebated Equipment)

<1> Yes
<3> No [goto 013]
<5> Don't know [goto 013]

Answer ===> [goto 011]

>011< Was the equipment purchased to replace old equipment or was it purchased as additional equipment?

<1> Purchased as replacement equipment
<2> Purchased as additional equipment [goto 013]
<3> Both replacement and additional
<5> Don't know [goto 013]

Answer ===>

>012< What type of fuel did the replaced equipment use?
<1> Gas
<3> Electric
<5> Don't know

Answer ===>

>013< Has the total amount of gas equipment changed since January 1, 1996? That is, have you increased or decreased the number of cooking equipment units fueled by gas since January 1, 1996.

<1> Increased
<2> Decreased
<3> Has not increased or decreased [goto 015]
<5> Don't know [goto 015]

Answer ===>

>014< On the whole, what would you say is the percent of change in your total gas cooking equipment since January, 1996.

<0-100>Percent
<999> Don't know

Answer ===>

SCG Commercial Equipment Replacement Evaluation Survey

We would like to know the number and types(s) of gas cooking equipment you presently use in your kitchen?

>15a1< Do you have.a Braising Pan(s)?

How many are in use now?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15a2< How many were in use prior to 1996?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15b1< How about Broilers / Cheesemelters?

How many are in use now?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15b2< How many were in use prior to 1996?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15c1< How about Cabinet Steamers?

How many are in use now?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

```
>15c2< How many were in use prior to 1996?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15d1<  How about fryers?

        How many are in use now?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15d2< How many were in use prior to 1996?

>15e1<  How about Griddles?

        How many are in use now?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15e2< How many were in use prior to 1996?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15f1<  How about Ovens?

        How many are in use now?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15f2< How many were in use prior to 1996?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>
```

```
>15g1<  How about Convection Ovens?

        How many are in use now?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15g2< How many were in use prior to 1996?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15h1<  How about Ranges?

        How many are in use now?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15h2< How many were in use prior to 1996?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15i1<  How about Steam Kettles?

        How many are in use now?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>

>15i2< How many were in use prior to 1996?
        <0>     none
        <1-10>  one to ten
        <11>    eleven or more
        <99>    Don't Know

        Answer ===>
```

>15j1< How about Steam Tables?

How many are in use now?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15j2< How many were in use prior to 1996?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15k1< How about Other Gas Cooking Equipment?

How many are in use now?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>15k2< How many were in use prior to 1996?
<0> none
<1-10> one to ten
<11> eleven or more
<99> Don't Know

Answer ===>

>016< Does this facility serve breakfast, lunch or dinner?
(Code each choice)

<1> Breakfast
<2> Lunch
<3> Dinner
<4> No meal served [goto 022a]
<9> Other meals, specify

Answer ===>

>016j< What percentage of the meals cooked at your facility are for breakfast?

<1-100> percent
<999> don't know

Answer ===>

>016k< What percentage of the meals cooked at your facility are for lunch?

<1-100> percent
<999> don't know

Answer ===>

>016l< What percentage of the meals cooked at your facility are for dinner?

<1-100> percent
<999> don't know
Answer ===>

>017< What is the number of meals cooked per week at this kitchen facility?

<0-10,000> meals per week

Answer ===>

>018< Has the number of meals cooked per week changed since January 1, 1996

<1> Yes, now have fewer
<2> Yes, now have more
<3> No [goto 021]
<5> Don't know [goto 021]

Answer ===>

>020< What was the approximate number of meals cooked per week before January 1, 1996

<0-20,000> meals per week
<9> Don't know
Answer ===>

Now, we would like to ask about the type and quantity of meals you serve and your hours and days of operation.

>021< What type of meals are served at this facility?
<1> Fast food/burgers, fries, tacos, etc.
<2> Pizza/sandwiches/bread, etc. baked on premises
<3> Self serve/cafeteria
<4> Table service
<5> Other

Answer ===>[specify]

>022a< How many days per week does the kitchen facility operate?

<1-7> days per week

Answer ===>

>022b< Does the kitchen facility operate all 12 months of the year?

<1> Yes
<2> No [goto 024]

Answer ===>

>023< Which months per year is this facility closed?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December

<99> None, we never close

Answer ===>

>024< Is the amount of cooking done this kitchen facility about the same throughout your business year?

<1> Yes [goto 026]
<2> No
<5> Don't know

Answer ===>

>025< During which months do you experience the heaviest volume?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> None, they are about the same

Answer ===>

Now, we would like to ask about any renovation you may have had on your building.

Building Changes Since Installation of the Measure

>026< Have you expanded, reduced, or had major renovations to your building space since January 1, 1995?(Check all that apply)

<1> Expanded space [gotto 026e]
<2> Reduced space [goto 026h]
<3> Renovated space [goto 028]

<4> No change in space used [goto 026d]

Answer ===>

>026e< How many square feet did you add during this expansion?

<100-100,000> square feet
<99> Don't know exactly

Answer ===>

>026f< What year did this expansion in occur?

<1> 1995
<2> 1996
<3> 1997

<9> Don't know [goto 028]

Answer ===>

>026g< What month did this expansion occur?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> Don't know

Answer ===>

>026h< How many square feet did you eliminate during this renovation?

<100-100,000> square feet
<99> Don't know exactly

Answer ===>

>026i< What year did this reduction in occur?

<1> 1995
<2> 1996
<3> 1997
<9> Don't know [goto 028]

Answer ===>

>026j< What month did this reduction occur?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> Don't know

Answer ===>

The last few questions are about non-cooking related equipment that you may have at your establishment.

Cooling System

>028< What type of fuel is used for the main cooling system at this facility?

<1> Gas
<2> Electric
<4> None, no air conditioning [goto 038]
<5> Other
<9> Don't know

Answer ===>

>029< What percentage of the total area of your establishment is air conditioned?

<1-100> Percent
<999> Don't know

Answer ===>

>030< Have you installed a new main cooling system since January 1, 1995?

<1> Yes [goto 031]

<3> No
<5> Don't know

Answer ===> [goto 038]

>031< Was this additional or replacement equipment?

<1> Additional
<2> Replacement [goto 032]

<5> Don't Know [goto038]

Answer ===> [goto 033]

>032< What type of fuel was used by the old cooling system?

<1> Gas
<3> Electric
<5> Other
<9> Don't know

Answer ===>

>033< What year was this main cooling system installed?

<1> 1995
<2> 1996
<3> 1997
<9> Don't know [goto 035]

Answer ===>

>034< What month was it installed?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> Don't know

Answer ===>

>35< Is this cooling system conditions the same amount of space as the
old cooling system?

<1> Yes, the same amount of space (square footage)
<2> No, cooling more space [goto 037]
<3> No, cooling less space [goto 037]
<9> Don't know

Answer ===> [goto 038]

>037< How much more or less space is the new cooling system conditioning?

<1-100> percent
<999> Don't know

Answer ===>

Heating System

>038< What percentage of the total area of your establishment is heated?

<0> none [goto 049]
<1-100> Percent
<999> Don't know

Answer ===>

>039< What type of fuel is used for the main heating system at this facility?

<1> Gas
<2> Electric
<3> Other
<9> Don't know

Answer ===>

>040< Have you installed a new main heating system since January 1 1995?

<1> Yes [goto 041]
<3> No
<5> Don't know

Answer ===> [goto 049]

>041< Was this additional or replacement equipment?

<1> Additional [goto 043]
<2> Replacement
<5> Don't Know

Answer ===>

>042< What type of fuel was used by the old heating system?

<1> Gas
<2> Electric
<3> Other
<5> Don't know

Answer ===>

>043< What year was this new heating system installed?

<1> 1995
<2> 1996
<3> 1997
<9> Don't know [goto 045]

Answer ===>

>044< What month was it installed?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> Don't know

Answer ===>

>045< What percentage of the total area of your building is affected by this new heating system?

<0-100>
<999> don't know [goto 049]

Answer ===>

>046< Is this heating system conditioning the same amount of space as the old heating system?
<1> Yes
<3> No, heating more space [goto 047]
<4> No, heating less space [goto 047]
<5> Don't know

Answer ===> [goto049]

>047< How much more or less space is the new heating system conditioning

<0-100> percent
<999> don't know

Answer ===>

Insulation

>0049< Do you know if the walls, ceiling, or pipes are insulated at your facility?

<1> Yes
<3> No [goto 055]
<5> Don't know [goto 055]

>050< Where is this insulation installed?

<1> Walls
<2> Ceiling
<3> Pipes
<9> Other location, specify

Answer ===>

>051< Has any new insulation been installed since January 1, 1995?

<1> Yes
<3> No [goto 055]
<5> Don't know [goto 055]

Answer ===>

>052< What year was it installed?

<1> 1995
<2> 1996
<3> 1997
<9> Don't know [goto 054]

Answer ===>

>053< What month was it installed?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> Don't know

Answer ===>

>054< What percentage of your facility is covered by this insulation?

<0> None
<1-100> Percent
<999> Don't know

Answer ===>

Water Heating Section

>055< What type of fuel does/do your current water heater(s) use? (Check all that apply)

<1> Gas
<2> Electric
<3> Other
<9> Don't know

Answer ===>

>056< How many water heaters do you have at your facility?

<0 -10> Water heaters
<99> Don't know

Answer ===>

>057< What is the capacity of the water heater(s)

<20-1000> gallon capacity
<9> Don't know

Answer ===>

>058< Have you installed a new water heater since January 1, 1995

<1> Yes
<3> No [goto 064]
<5> Don't know [goto 064]

Answer ===>

>059< What type of fuel did your old water heater use?

<1> Gas
<2> Electric
<3> Other
<5> Don't know

Answer ===>

>060< What year was your new water heater installed?

<1> 1995
<2> 1996
<3> 1997
<9> Don't know [got 062]

Answer ===>

>061< What month was it installed?

<1> January
<2> February
<3> March
<4> April
<5> May
<6> June
<7> July
<8> August
<9> September
<10> October
<11> November
<12> December
<99> Don't know

Answer ===>

>062< Was this new water heater an addition or did you replace the existing equipment?

<1> Additional equipment
<2> Replacement
<5> Don't know [goto 064]

Answer ===>

>063< What percentage of hot water is heated by this new equipment?

<1 -100> Percent
<999>

Answer ===>

>064< Has the amount of dishwashing changed since January 1, 1996?

<1> Yes [goto 065]
<3> No
<5> Don't know

Answer ===> [goto 066]

>065< Has the amount of dishwashing increased or decreased since the January 1, 1996?

<1> Increased
<3> Decreased
<5> Don't know

Answer ===>

>066< Is natural gas used for purposes other than space heating or cooling, water heating or cooking at your establishment?

<1> Yes, specify
<3> No
<5> Don't know

Answer ===>

Thank you very much for participating in our survey.