A UERG Program in Analysis of California Energy Consumption Data Report

UER-189

Program in Analysis of California Energy Consumption Data

First Year Project Results

Joseph H. Eto
Program Coordinator

April 1988

This research was sponsored by the Universitywide Energy Research Group with funding from the California Energy Commission.

Contents

Preface

Accurate data on end-use energy demands has become increasingly important for forecasting as utilities integrate demand-side programs into their resource plans. For California's residential sector, a major source of this information has been provided by mail surveys administered by the major utilities.

To assess the opportunities for improvements in residential survey design, data collection and analysis, and, ultimately, demand forecasts, the California Energy Commission (CEC) has sponsored five research projects at four University of California campuses through the Universitywide Energy Research Group (UERG). This volume contains summaries of the first year efforts of these projects. The summaries have been assembled to introduce major findings and policy recommendations in preparation for a jointly sponsored CEC/UERG conference for CEC and California utility staff members.

In the first paper, Dane Archer, Elliot Aronson, and Tom Pettigrew of the Sociology and Psychology departments at U.C. Santa Cruz present an analysis of sampling and survey or unit non-response biases in the 1983 Los Angeles Department of Water and Power (LADWP) and 1983 San Diego Gas and Electric (SDG&E) residential surveys. In the second paper, Trudy Ann Cameron of the Economics department at U.C. Los Angeles develops an innovative econometric technique to evaluate participants in utility-sponsored residential retrofit programs in the LADWP service territory. In the third paper, C.W.J. Granger and Halbert White of the Economics department at U.C. San Diego analyze several SDG&E residential surveys to estimate structural changes in unit energy consumption (UECs) over time. In the fourth paper, Bruce Hackett, Howard Schutz, and Loren Lutzenhiser of the Sociology department at U.C. Davis use a SDG&E residential survey and Census data to evaluate the benefits of disaggregating residential energy forecasts along social dimensions. In the fifth paper, Paul Ong of the Architecture and Landscape Planning department at U.C. Los Angeles, and Nirvikar Singh and Suzanne Holt of the Economics department at U.C. Santa Cruz study methods to correct for question non-response using a PG&E residential survey and subsequent audit information.

The Program in Analysis of California Energy Consumption Data was managed by Tom Gorin for the California Energy Commission. The first year program coordinator for the Universitywide Energy Research Group was Carl Blumstein. Valuable direction for the Program was provided throughout the course of research by Mike Jaske, Chief Energy Forecaster for the California Energy Commission.

ASSESSING AND MINIMIZING BIAS IN THE RESIDENTIAL APPLIANCE SATURATION SURVEYS

Prepared for:

PROGRAM IN ANALYSIS OF CALIFORNIA ENERGY CONSUMPTION

of the
California Energy Commission
and the
Universitywide Energy Research Group (UERG),
University of California

April 6, 1988

Energy Conservation Research Group
Stevenson College
University of California
Santa Cruz, CA 95064

Principal Investigators

Dane Archer, Professor of Sociology
Elliot Aronson, Professor of Psychology
Thomas F. Pettigrew, Professor of Social Psychology

Research Associates

Scott Coltrane
Marti Hope Gonzales
Marty Lucia

This paper summarizes findings from a larger study conducted for the Program in Analysis of California Energy Consumption of the California Energy Commission (CEC) and the Universitywide Energy Research Group (UERG), University of California. In the first section of the paper, we highlight the major findings of the larger study, followed by a brief review of common sources of bias in sample surveys. We then present some general findings from our analysis of two Residential Appliance Saturation Surveys (RASS), and conclude with a summary of recommendations for minimizing biases in future RASS.

SUMMARY OF FINDINGS

This study examines the importance of response rates for RASS surveys. When there are large and systematic nonresponse biases, survey results cannot be confidently generalized to the population of interest. It is important, then, to test for the possibility of such biases in RASS surveys.

The CEC staff provided two 1983 RASS surveys for analysis - the Los Angeles Department of Water and Power RASS (LADWP) and the San Diego Gas and Electric RASS (SDG&E). Our study demonstrates serious nonresponse biases in both surveys and recommends a series of changes for future RASS surveys to alleviate the problem.

The 1983 LADWP RASS contains two large sampling biases. The first involves a sampling frame bias introduced by excluding all customers with less than six months of valid electric meter readings. This exclusion eliminated 300,000 customers. The second bias emerged from the low total response rate of 33.3% - far below the standard of 75 to 80%. Both of these biases tend to exclude small energy users. When combined, they caused a sharp overrepresentation of higher socio-economic status respondents who tended to use more electricity, to have more appliances, and to have taken more conservation measures than other customers. We estimate that these biases led to constant errors in excess of 50% for selected appliances.

Are such biases shared by other RASS surveys? To find out, we analyzed the

1983 SDG&E RASS. While no sampling frame bias was uncovered, this survey, too, suffered from systematic nonresponse bias. It obtained only a 51% total response rate; and the biases thus introduced were similar to those uncovered in the LADWP study. Lower-status customers were significantly less likely to have responded. Hence, once again energy usage, appliance ownership, and conservation measures tend to be overestimated by the SDG&E RASS results. While not quite as extensive as the LADWP biases, these SDG&E problems could, for selected dependent variables, approach an error rate of 50%.

We conducted an extensive review of the research literature to see if there are effective techniques for correcting for nonresponse biases after the fact. Our conclusion is that such "patching" methods are at best crude approximations; and, indeed, they cannot be generally applied to past RASS data because the data needed for their use have generally not been recorded. This review leads us to a set of recommendations for future RASS studies that would allow the utilities and the CEC to monitor more carefully and correct for similar biases in the future.

In the concluding chapter of our final report, we offer a complete set of recommendations for a new RASS survey design - one that would sharply alleviate nonresponse problems and allow more confident generalizations in the future. The essence of these recommendations revolves around a different survey philosophy that emphasizes far smaller sample sizes and far higher response rates. No greater costs would be involved; indeed, there might well be modest cost savings. The need to enlarge the size of particular subsets of households can be efficiently handled by wider use of refined stratification and oversampling. But the key goal is to pursue reluctant respondents. Toward that end, we recommend three mail waves (rather than the present two) in addition to the final phone wave and the use of Dillman's Total Design Method (TDM). The larger report provides details for these recommended procedures together with an extensive bibliography that supports them.

SURVEY BIASES

Problems of nonresponse and other biases in surveys are typically treated like the weather. Everyone who depends on survey data knows at least vaguely about survey biases, worries about them on occasion, and hopes for the best. But rarely does anyone do anything about them. Indeed, survey users often simply ignore, on critical occasions, the error introduced by survey biases. They may even come to regard survey biases as "benign" when there are actually serious problems.

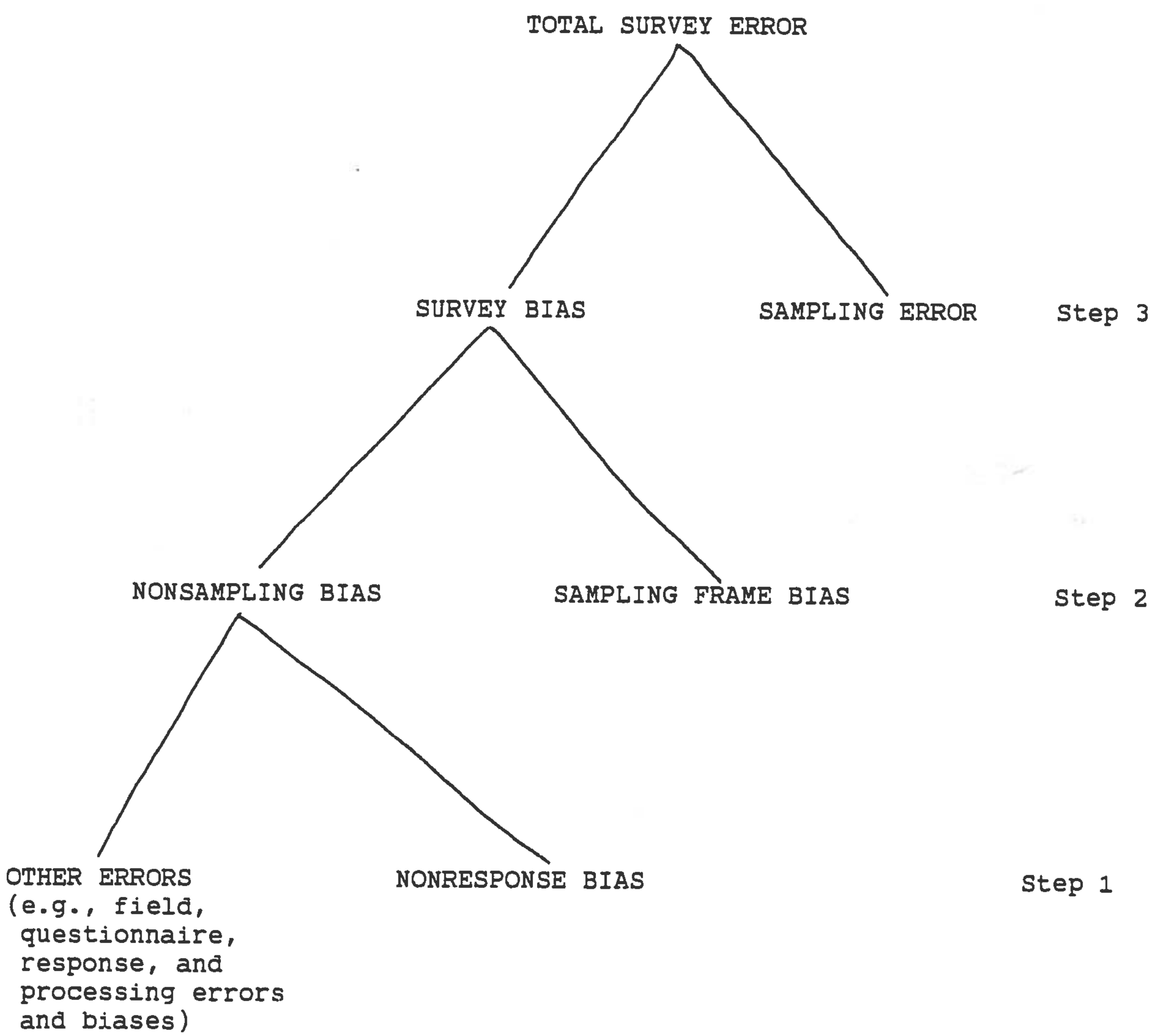
The intent of any survey is to represent accurately the population from which it samples--that is, the sample should closely resemble characteristics of the population. Otherwise, survey findings cannot be generalized beyond the sample and back to the population of interest.

Important issues in survey methodology will be discussed in the following sections: sampling error, sampling frame bias, sample size and selection, response rate, and nonsampling bias (including nonresponse and other errors). Discussion is limited to the most important of the errors and biases inherent or potential in all sample surveys.

Figure 1 and the corresponding Glossary provide a convenient summary of "Total Survey Error." The flowchart in Figure 1 indicates how a survey is subject not only to survey error, but also a variety of survey biases. Note the downward branching from total survey error. The end branches trailing off to the right signify the main steps which need to be assessed and, if possible, neutralized. The survey researcher should start at Step 1, nonresponse bias, making sure that the response rate is sufficiently high enough to eliminate this as a possible significant bias.

The researcher can then progress upwards to Step 2, sampling frame bias, making sure that the sample selected is a probability sample in which there are no systematic differences between the universe and the sampling frame. Assuming that serious sampling frame and nonresponse biases are not present, the researcher can then compute the sampling error and assess the survey's precision. It is a major theme in this paper that surveys can be invalidated by <u>any</u> of a series of errors and biases and that, contrary to lay conceptions about survey research, sampling error is only <u>one</u> of these.

Figure 1 Assessing Total Survey Error.



(Refer to Glossary of Survey Terms for Definitions.)

Glossary of Survey Terms

1.) Total Survey Error = The sum of errors of all types that affect the accuracy (i.e., difference between a sample statistic and a corresponding population parameter) of a survey.

2.) Sampling Error = Expressed as the standard error, sampling error is an estimate of the variability of the sample mean around the population mean.

3.) Survey Bias = That part of total survey error which includes all biases from samples (sampling frame bias) as well as in complete censuses (nonresponse bias).

4.) Sampling Frame Bias = Sample selection procedures which fail to provide each population element (person) with an equal chance for being selected in a probability sample.

5.) Nonsampling Bias = Factors resulting apart from the accuracy of the original sample--most notable of the nonsampling biases is nonresponse bias--i.e., failure of sampled individuals to complete part or all of a survey.

6.) Nonresponse Bias = Occurs when (1) there is a failure to obtain observations on some members of the sample, and (2) there are systematic differences between respondents and nonrespondents.

7.) Other Errors = A catch-all grouping, referring to the difference between the true and the observed measurements. The difference may be systematic (bias) or random (error). Examples of such errors and biases include, but are not limited to: poorly worded questions, biased response patterns, and mechanical or other data entry problems.

Sampling Error

Expressed as the standard error, sampling error refers to the variability of the population estimates around the population mean from repeated samples. With increasingly large samples, this variability decreases. But, as we shall note later in recommending smaller, more precisely drawn samples, the decrease in the size of the sampling error with larger samples rapidly diminishes as the sample size rises. Thus, far smaller samples than those typically drawn in past RASS surveys will still retain appropriately small sampling error.

Two additional and related points concerning sampling error need to be stressed. Strictly speaking, sampling error is not "sampling bias." Sampling error is merely variable error that tends to reduce in magnitude as the sample size reaches sufficient numbers. Sampling bias involves constant error; and it does not decrease as the sample size increases (Yates, 1981). Yet there is often a confusion between sampling error and sampling bias.

We have noted claims in several RASS reports of minimal sampling errors due to large sample sizes as if this removed all problems of survey bias. For example, citing confidence limits of 3% "error" can provide misleading assurance to an unwary reader. It would apply in practice when representative sampling and a 100% response rate had been obtained. But when a more typical RASS response rate of 55% applies, combined with occasional sampling frame bias, the <u>true total error rate</u> is far greater than 3% because the sampling frame bias and nonresponse bias may be orders of magnitude larger than the sampling error. The reader will recall from Figure 1 that the total survey error not only involves sampling error (which is usually reported) but also nonresponse bias, sampling frame bias, and other errors (which are rarely reported and calculated as part of the total survey error). Our report provides concrete examples of the debilitating effects that nonresponse bias and sampling frame bias have on total survey error in RASS.

This critical distinction between sampling error and nonresponse bias is a basic differentiation that is critical to the recommendations of our study. In brief, we believe that RASS surveys have in the past focused too much attention on reducing sampling error with huge samples at the expense of nonresponse bias. In the final chapter of our report, we propose a set of recommendations for future RASS that redress this imbalance.

Sampling Frame Bias

Bias, as defined by the sampling expert Leslie Kish, "refers to systematic errors that affect any sample taken under a specified survey design with the same constant error" (1965, p. 509). While sampling error decreases as the sample size approaches the population, sample bias remains as constant error which does not decrease with larger sample sizes.

Most troublesome of the sampling biases is sampling frame bias. The sampling frame includes the procedures to account for all the elements in the universe as well as an actual listing. A perfect frame is one that includes every element once and only once. Sampling frame bias occurs when there are systematic differences between the universe and the sampling frame. In the present case of RASS surveys, sampling frame bias would arise when there is a systematic difference in energy usage between the population of a utility's households and the sampling frame from which the survey sample was drawn.

In the larger report we note in detail precisely such a sampling frame bias. The LADWP RASS of 1983 excluded from its sampling frame all customers without valid electric meter readings for six months. This decision eliminated over 300,000 customers from the sampled universe. Though we are unable to test it directly, there are good reasons to believe that the excluded customers tended to be considerably smaller users of electricity than those retained in the sampling frame.

Sample Size and Selection

One common misconception holds that increasing the size of the sample in order to acquire more responses to the survey reduces the problem of nonresponse bias. On the contrary, "Once the planned sample is large, say equal to 1,200, increasing the size of the sample even dramatically does not alleviate nonresponse bias" (Wayne, 1975-1976, p. 559, emphasis in original). It is critical to realize that the total survey error (which includes sampling error, sampling frame bias, and nonresponse bias) is calculated from the total sample, not just from the respondents.

Generally, the size of the sample will depend on the degree of accuracy required for the sample (sampling error) and the extent to which there is variation in the proportion of responses for important questions asked. There are diminishing returns for increasing sample sizes, since beyond a certain point the extra precision gained by increasing the sample size does not outweigh the additional costs involved. Many survey companies limit their samples to 2000 since the cost of generating larger samples has insufficient payoff in terms of accuracy (de Vaus, 1986).

As is well known, the fundamental property of probability samples is that each unit has an equal chance of selection. Or, if different sampling ratios have been used, these ratios are known and the data can be weighted to retrieve the population estimates just as if each unit had an initial equal probability of selection.

This straightforward handling of probability sampling was on occasion violated in the RASS studies we reviewed. In the LADWP RASS of 1983, for instance, the sampling frame bias just cited was compounded by the drawing of a non-probability sample. Thus, the number of customers selected from each sampling cell remained at the original one per cent target in spite of the biasing fact that different proportions of customers had already been removed from the sampling cells by the length of residence selection. In other words, the LADWP RASS of 1983 had its sample drawn as if the sampling frame bias of removing new customers had not taken place. Consequently, longer-term customers retained in the sampling frame who lived in highly transient areas where a large proportion of new customers had been screened out had a far larger probability of selection than other longer-term customers living in stable areas with relatively few new customers screened out.

Stratified Sampling

One common misunderstanding about probability samples is to regard them as "representative" - that is, that such samples will accurately reflect the proportions of various subgroups within the population. But if the sampling design did not originally attend to these subgroups, sampling error may produce proportions widely divergent from their true proportions - especially for relatively small subgroups. The effective way to handle this problem is to stratify for the subgroups of particular interest in the original sampling design.

Stratification essentially involves multiple probability samples - a separate one for each subgroup for each subgroup cell (e.g., heavy users in a particular climate zone) - <u>with the cell sizes set to their known ratios in the universe</u> (<u>population</u>).

The total sample size is often dictated by the sample size needed for the smallest subgroup - though special oversampling can be done of important subgroups. According to Hoinville et al. (1978), the smallest subgroup should have between 50 and 100 members. In most surveys which explore variations within the population, samples of 1,000 to 5,000 are used, far fewer than the 15,000 often sampled in the RASS surveys.

We observed that the two RASS surveys we examined did stratify for key variables of interest. However, a constant one per cent sample (LADWP) and a constant two per cent sample (SDG&E) was maintained for each strata, presumably to maintain large enough cases for the smallest of strata (such as individually metered multi-family dwellings in the Maritime Zone with electric space heating who consume large amounts of energy). Greater use of disproportionate stratified sampling designs are advocated in our final report. This technique offers one answer (among others) to an objection we anticipate to our major recommendation for future RASS surveys to use far smaller samples in order to achieve much higher response rates. Stratification, together with oversampling for important subsets, offer a cost-saving means of achieving this goal without the necessity of using enormous sample sizes.

While a larger sample size may indeed acquire more responses overall, the smaller sample size allows for a more intense last wave (phone or personal interview) to be given to the nonrespondents. This will decrease the non-response rate which then leads to a more confident analysis of the data.

Survey Response Rate

This rate simply calculates the percentage of the original sample drawn from whom at least a partial set of responses is obtained. There are several types of nonresponders who often differ somewhat from each other: not-at-homes (often single adults in large metropolitan areas; and blacks more than whites); outright refusals (often married couples and older sample members in large metropolitan areas); and a miscellaneous category of great diversity (e.g., no fixed address or phone; illiterates; non-English speakers; the institutionalized; hence, this category is heavily overrepresented among the poor, single, and elderly in large metropolitan areas) (Dunkelberg & Day, 1973; Mayer, 1964; Smith, 1983). Note how the poorer and less educated - the lower socio-economic status (SES) - sample members in large urban areas are highly likely to be among the unsurveyed in all three of these categories.

This melange of types introduces a variety of different, and sometimes conflicting, biases. For some kinds of variables, these biases can prove either irrelevant or they largely cancel themselves out. If there are no constant, systematic differences between the respondents and nonrespondents on the key variables of interest, then no direct bias is introduced by the nonrespondents. But, as our analysis reveals, the RASS surveys are not so fortunate.

The important point, however, is that response rates do not in themselves measure survey bias even though they are highly and negatively related to the potentiality of such biases. A quite high response rate, say 90%, largely eliminates the possibility of serious bias. While rates typical of the RASS surveys - 35% to 60% - raise the possibility of extremely serious biases. In short, serious nonresponse bias requires two conditions: a relatively low response rate (below the 75% - 80% sometimes required on federal contracts) <u>AND</u> some systematic difference between the respondents and nonrespondents on the item of interest (Fowler, 1984).

Often differential tests are made using demographic variables that are known for the nonrespondents as indirect indicators of bias: Are the proportion of men, college-educated, farmers, those over 65, etc., of the respondents significantly different from the known proportions of the sampled universe? When the proportions are similar, it is sometimes claimed that there is no non-response bias. But such claims overlook that the demographic tests are only indirect indicators at best of any response bias on the non-demographic items of interest. There may well be (and often are) critical differences between the respondents and nonrespondents on the major variables of the study even when their demographic distributions closely resemble each other. One strength of our study is that we were able to analyze direct information on the nonrespondents on a key variable of importance - their energy usage.

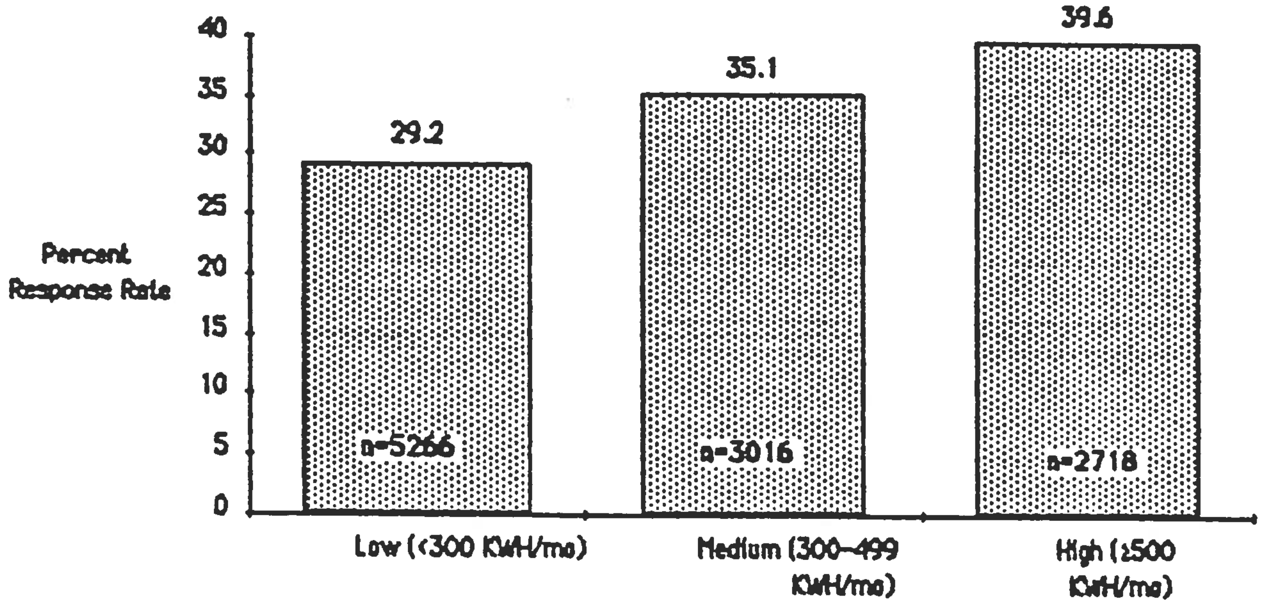
Virtually every study of the patterns of return for self-administered questionnaires shows that early returns are biased. Samples of data resulting from returns of 50% or less almost always depart from the total sampled population. This is especially true for data of the type collected in the RASS surveys, since electrical consumption, appliance ownership, and conservation are directly related to socio-economic status (SES). SES, in turn, is one of the best predictors of the propensity to return mail questionnaires.

Electrical Consumption

Households that failed to respond to the first or second mailing of the the 1983 LADWP RASS used significantly less electricity than those who returned the mailed questionnaire ($t=6.7$, $p<.001$, two-tailed test). Similarly, households not interviewed in the phone/onsite follow-up used less electricity than those interviewed ($t=1.7$, $p<.05$, one-tailed test). As shown in Table 1, those households in the lowest use category (<300 kWh/mo.) had the lowest overall response rate (29.2%). Households with medium use (300-499 kWh/mo.) had the next highest response rate (35.1%), and higher users (>500 kWh/mo.) had a higher rate of response (39.6%). These differences are statistically significant (Chi-square = 58.8, $p > .001$). When high users were further disaggregated, those using over 1000 kWh/month were found to be the most likely to return the questionnaires, with a response rate of 42.6%. These differential response rates suggest that RASS data are representative of higher-volume electrical consumers.

Table 1.

Response Rate by Electrical Consumption: 1983 LADWP RASS (*)



*chi-square=58.8, p<.001

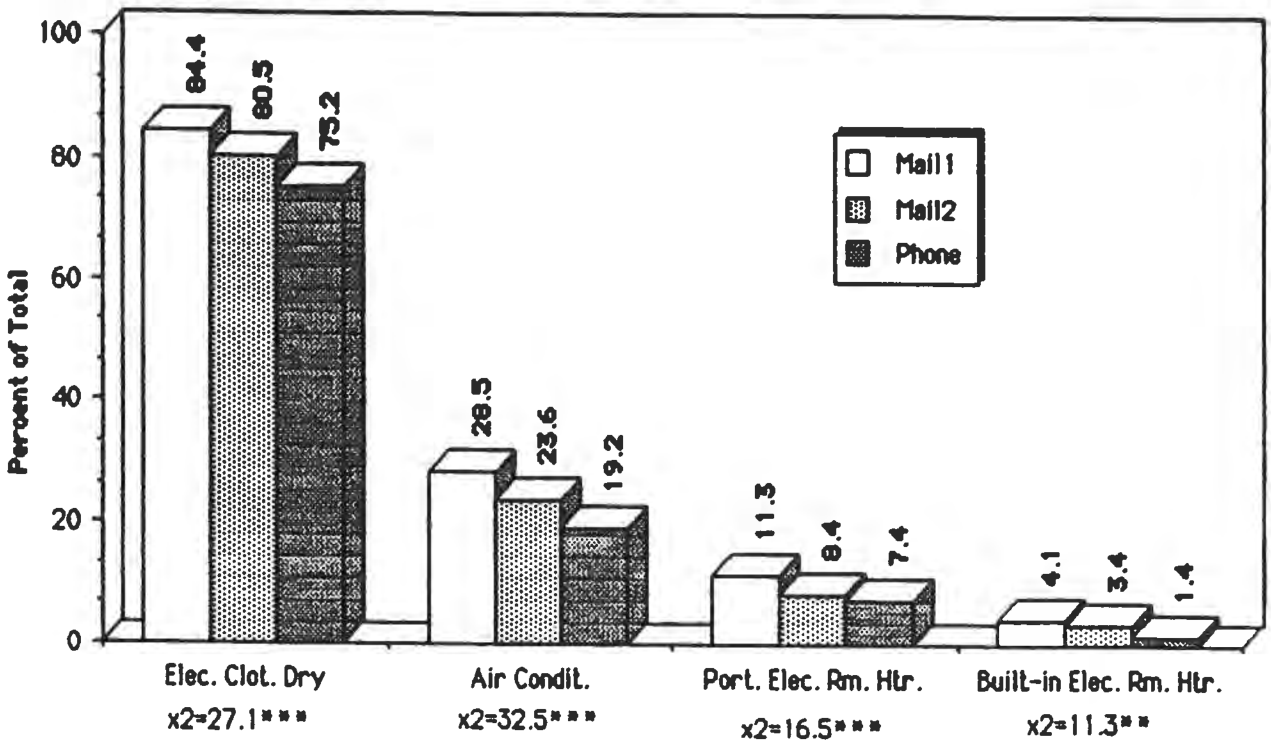
Aggregate data from the 1983 San Diego RASS also show that low-volume electrical consumers were the least likely to return the questionnaires. Since consumption data on individual nonrespondent households were not available for San Diego, we were unable to replicate our L.A. findings. Nevertheless, many of the same patterns of appliance holdings and conservation actions found in Los Angeles were found in San Diego. In general, those who initially respond to the RASS survey use more electricity, own more applicances, and install more conservation devices than those who respond later.

Appliance Characteristics of Early vs. Late Respondents

The earlier a household responded to the the 1983 San Diego RASS, the more likely that household was to own major electricity consuming appliances. Table 2 shows a statistically significant linear relationship between response wave (1st mail, 2nd mail, phone) and the presence of selected appliances including clothes dryers, air conditioners, portable electric room heaters, and built-in electric room heaters. The figures in Table 2 apply to the simple presence of air conditioning or clothes dryers in a residential complex. Equally consistent findings obtain, however, for private ownership of these appliances. (Air conditioning, Chi-square = 29.9, $p<.001$; clothes dryers, Chi-square = 33.1, $p<.001$).

Disaggregating by housing type shows that the trends evident in Table 2 hold for both single family and multi-family dwellings. Early respondents living in single family homes are significantly more likely to have electric clothes dryers (Chi-square = 17.4, $p<.001$), air conditioning (Chi-square = 21.0, $p<.001$), and either portable or electric room heaters (Chi-square = 15.7, $p<.001$). Similarly, early respondents living in multi-family housing are significantly more likely to have electric clothes dryers (Chi-square = 11.0, $p<.004$), air conditioning (Chi-square = 14.7, $p<.001$), and either portable or electric room heaters (Chi-square = 7.3, $p<.02$). Relying on a RASS return sample that is predominantly composed of early returns (in 1983 San Diego RASS:

Table 2. Appliance Characteristics of Early vs. Late Respondents: 1983 SDG&E RASS



72%=1st mail response, 20%=2nd mail response) overestimates the presence of these and other appliances in the population.

Conservation Action

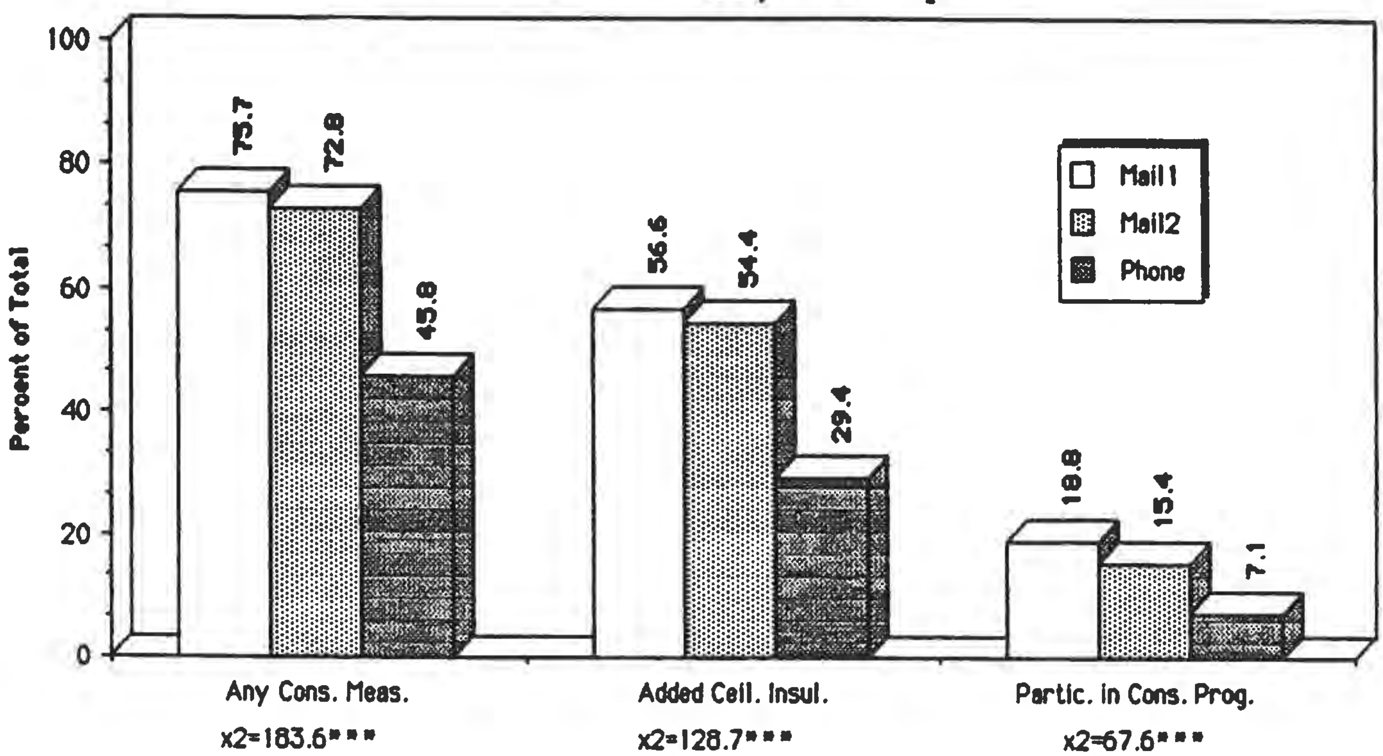
The relationship between conservation action and timing or difficulty of response is even more pronounced, as evidenced in Table 3. Households returning the 1983 SDG&E RASS questionnaire in the first mail wave were the most likely to have undertaken some form of conservation, with 2nd wave mail respondents slightly less likely to have undertaken conservation, and phone respondents appreciably less likely to have done so (Chi-square = 183.6, p<.001). [1]

Two examples of specific conservation actions are also included in Table 3. Adding ceiling insulation, the single most efficient conservation retrofit, was most often undertaken by early returners. Over 50% of households returning the mailed version of the RASS had installed ceiling insulation, compared to less than 30% of RASS phone respondents (Chi-square = 128.7, p<.001). Similarly, early returners were twice as likely to have participated in utility and government sponsored conservation programs, including audits, loans, and tax credits (Chi-square = 67.6, p<.001).

As with appliance holdings, early respondents were more likely than late respondents to have taken conservation action, regardless of whether they lived in single family or multi-family housing. Eighty percent of first or second wave mail respondents in single family homes reported taking conservation action, whereas 54% of phone respondents in single family homes reported doing so (Chi-square = 110, p<.001). Fifty-eight percent of multi-family dwelling

[1] Conservation measures included in this RASS question were: (1) adding ceiling insulation; (2) installing a water heater blanket; (3) installing a clock/setback thermostat; (4) installing a showerhead flow restricter; (5) replacing lightbulbs with lower wattage or fluorescent bulbs; (6) putting in caulking or weatherstripping around windows or doors; (7) putting in insulation around water pipes or air ducts; (8) installing a swimming pool cover; (9) adding wall insulation; (10) adding an attic turbine ventilator; (11) adding window film, glazing, or external shading; or (12) reducing the temperature setting on hot water heater.

Table 3. Conservation Characteristics of Early vs. Late Respondents: 1983 SDG&E RASS



mail respondents had taken conservation action, whereas 28% of multi-family dwelling phone respondents had done so (Chi-square = 47, p<.001). Similar results are found when disaggregating by housing tenure. Early respondents are significantly more likely to have taken conservation action regardless of whether they are homeowners or renters.

In addition to asking whether residents had installed ceiling insulation, RASS question No. 32 asked if the residence had ceiling insulation. Similar to the results reported above, early respondents were significantly more likely to live in dwellings with ceiling insulation present. 74% of first wave respondents reported that their dwelling had ceiling insulation, compared to 69% of 2nd wave respondents, and 65% of 3rd wave (phone) respondents (Chi-square = 30.9, p<.001).

Early responding households do not represent the conservation actions of later responding households. We estimate that nonrespondent households are even less likely to have taken conservation actions than phone respondents, suggesting a serious underestimation of future conservation potential based on RASS data. This disturbing finding raises the possibility of prematurely concluding that future conservation potential is more limited than it actually is. Basing estimates of conservation saturations on RASS returns is unwarranted. Until higher response rates are obtained, we cannot accurately predict appliance saturation, nor the extent of past or potential conservation program utilization.

Overall Bias in Residential Appliance Saturation Surveys

For the LADWP RASS we find that the two major forms of bias uncovered by our analysis are in the same direction: sampling frame decisions and overall nonresponse tended to result in estimates of the "average" household that were actually representative of households with higher-than-average socio-economic-status (as measured by Census data on income, education, and homeownership). This produced RASS survey results that overestimated households with higher electrical usage, more appliance holdings, and more conservation installations.

While sampling frame biases for the 1983 San Diego RASS are not as severe as for the 1983 LADWP RASS, the nonresponse biases are equally significant. In spite of higher overall response rates, the SDG&E RASS data are not representative of the population they are intended to represent. While conflicting biases lessen some of the impacts of the nonresponse problem for San Diego, marked differences remain. In general, the San Diego biases are similar to those found in Los Angeles, with lower socio-economic-status households less likely to respond to the RASS survey. Again, this produced RASS data that overrepresented moderately high-volume customers, those with more appliances, and those who had installed conservation devices.

While we cannot precisely estimate the impact of each form of bias from the available data, we conclude that the sampling frame biases and nonresponse biases combine to produce substantial survey bias. This bias probably exceeds 50% for selected appliances, and far outweighs the normal sampling error inherent in all survey research. Our conclusion, summarized below and presented in our final report, is that efforts should be directed toward the collection of more valid data, rather than trying to correct for bias by adjusting the severely flawed 1983 data.

A DIFFERENT SURVEY PHILOSOPHY

We believe that more efficient survey methods could improve the forecasting potential of the RASS surveys. On the following pages we present suggestions for conducting future RASS surveys that will substantially increase the validity of results and will allow for more accurate estimates of energy consumption and conservation potential in California.

Our recommendations are predicated on a survey philosophy that is somewhat different from that expressed in documents describing sample selection criteria for past RASS surveys (e.g., McCarthy, 1981). Our strategy entails the vigorous pursuit of relatively few subjects and is based on a philosophy that stresses the importance of eliminating systematic biases, especially nonresponse and sampling frame bias. Previous RASS surveys have utilized extremely large samples, based on a philosophy that stressed the importance of minimizing standard errors as computed from statistical formulas.

Following the accumulated wisdom of five decades of research on survey bias, we strongly recommend a leaner research design that entails the serious pursuit of reluctant respondents. The only way to limit nonresponse bias is to increase the response rate. Consequently, we recommend a variety of procedures to ensure response rates in the 70% to 80% range. These include three mail waves, postcard reminders, and an intensive effort on the fourth wave telephone interviews to reach reluctant respondents. By limiting initial sample sizes, we suggest that the rigorous procedures we recommend for ensuring high response rates will be no more costly than current RASS. The procedures we recommend are practical and well-tested, and have the potential for producing unbiased estimates of appliance saturations throughout California.

SUMMARY OF RECOMMENDATIONS

In this section we summarize some of our major recommendations for minimizing RASS bias. The first three suggestions are taken from Chapter 4 of the final report, and address the problem of nonresponse:

1. External population checks should be routinely made for each RASS between the respondents and both the nonrespondents and the population. These checks should be run on all relevant and available demographic variables, such as SES and urban-rural, as well as on energy usage as determined from bills. This assumes that variables and their values have been standardized according to Census format.

2. The final phone wave needs to be more fully utilized to help assess each RASS nonresponse bias by: (a) routinely keeping a data sheet on each nonrespon-

dent, including data on why the interviewer was not secured, times of the day called, primary home language, etc.; (b) for respondents, additional questions should be added to the mail questionnaire, including the number of times called and some availability at home information beyond that already tapped in the mail data; (c) calls to those who initially refused (either in the mail or phone waves) made by a specialist and with extra data gathered on those who do grant interviews in order to assess refusals' characteristics.

3. An additional probability sample should be drawn for each RASS of the nonrespondents of the previous RASS who reside at the same address as previously. As suggested by Kish, these data could then be used as replacements for similar nonrespondents in the current RASS.

The following steps summarize Dillman's (1978) Total Design Method presented in section 5.2 of the final report. These guidelines represent our study's more important and practical recommendations for designing and implementing high quality surveys.

TOTAL DESIGN METHOD (TDM)

WEEK 0. The Original Mailing -- FIRST WAVE
* sent first-class, on a Tuesday
* with a carefully composed cover letter.

WEEK 1. Postcard Reminder
* thanks those who already responded and
* reminds those who have not.

WEEK 3. Letter and Replacement Questionnaire -- SECOND WAVE
* sent first-class to all nonrespondents,
* with a shorter and more insistent cover letter.

WEEK 7. Letter and Replacement Questionnaire -- THIRD WAVE
* this time sent by certified mail to all nonrespondents.

WEEK 9. Letter Reminder
* sent to a one-third sample of the nonrespondents, to
* thank those who already responded,
* remind those who have not responded, and
* alert the nonrespondents to a forthcoming phone interview.

WEEK 10. Phone Survey -- FOURTH WAVE
* a minimum of seven recalls for the hard-to-reach,
* all refusers to be given one last call by a specialist,
* utilize the Kish replacement interviews, and
* collect additional data to assess nonresponse bias
* (e.g., number of times called, home language).

The reader should note that there are other critical issues related to sampling not included in the above list. As emphasized in the final report, we recommend:

1. Total sample sizes designed in the 1,000 to 1,500 range, which maximizes resources for more efficient follow-ups of the temporary nonrespondents. The final response rate should be in the 70% to 80% range for RASS.

2. A careful review of the rationale for specific stratification designs. If it is determined that extremely small strata must be maintained, we suggest that these strata be oversampled, with results weighted back to the population at large to represent their true proportion.

3. Results be combined over several surveys, or focus on selected subgroups in alternate surveys, in order to enlarge special subsets of customers efficiently.

4. 4. The use of a general package of techniques (as explained in the final report, section 5.2) in the design stage, including saliency, sponsorship, and incentives which can dramatically increase response rate.

5. To facilitate the analysis of future RASS, additional recommendations are presented in the final report concerning the collection, standardization, provision, and retention of data.

SELECTED REPORT REFERENCES

Alwin, D. (1977). Making errors in surveys: An overview. Sociological Methods and Research, 6, 131-150.

Anderson, R., Kasper J., & Frankel, M., & Associates. (1979). Total survey error: Applications to improve health surveys. NORC Series in Social Research. San Francisco: Jossey-Bass.

Andreason, A. (1970). Personalizing mail questionnaire correspondence. Public Opinion Quarterly, 34, 237-277.

Armstrong, J., & Overton, T. (1977). Estimating nonresponse bias in mail surveys. Journal of Marketing Research, 14, 396-402.

Baumgartner, R., & Heberlein, T. (1984). Recent research on mailed questionnaire response rates. In D. Lockhart (Ed.), Making effective use of mailed questionnaires. San Francisco: Jossey-Bass.

Blau, P. (1964). Exchange and power in social life. New York: Wiley.

Bryson, M. (1976). The Literary Digest poll: Making of a statistical myth. The American Statistician, 30, 184-185.

Carpenter, E. (1974). Personalizing mail surveys: A replication and reassessment. Public Opinion Quarterly, 38, 614-620.

Crossley, H., & Fink, A. (1951). Response and nonresponse in a probability sample. International Journal of Opinion and Attitude Research, 5, 1-19.

Daniel, W. (1975). Nonresponse in sociological surveys: A review of some methods for handling the problem. Sociological Methods and Research, 3, 291-307.

De Leeuw, E., & Hox, J. (1985). Recente ontwikkelingen in schriftelijk onderzoek. [Recent developments in mail research.] Jaarboek Van De Nederlandse Vereniging Voor Marktonderzoekers. [Yearbook of the Dutch Association for Market Researchers.] Haarlem, The Netherlands: De Vrieseborch.

De Leeuw, E., & Hox, J. (in press). The effect of an advance letter, a reminder by telephone, and the enclosure of a second questionnaire on the response to mail surveys. Kwantitatieve Methoden.

RASS BIAS REFERENCES

DeMaio, T. (1980). Refusals: Who, where, and why. Public Opinion Quarterly, 44, 223-233.

de Vaus, D. (1986). Surveys in social research. Contemporary Social Research Series, 11. London: George Allen & Unwin.

Dillman, D. (1978). Mail and telephone surveys: The total design method. New York: Wiley-Interscience.

Dillman, D., Frey, J. (1974). Contribution of personalization to mail questionnaire response as an element of a previously tested method. Journal of Applied Psychology, 59, 297-301.

Dillman, D., Dillman, J., & Makela, C. (1984). The importance of adhering to details of the total design method (TDM) for mail surveys. In D. Lockhart (Ed.), Making effective use of mailed questionnaires. San Francisco: Jossey-Bass.

Dunkelberg, W., & Day, G. (1973). Nonresponse bias and callbacks in sample surveys. Journal of Marketing Research, 10, 160-168.

Durbin, J. & Stuart, A. (1954). Callbacks and clustering in sample surveys. Journal of the Royal Statistical Society, 117, 387-428.

Eicher, D., & Habermehl, W. (1981). Predicting response rates to mailed questionnaires. American Sociological Review, 46, 361-363.

El-Badry, M. (1956). A sampling procedure for mailed questionnaires. Journal of the American Statistical Association, 51, 209-227.

Ellis, R., Endo, C., & Armer, J. (1970). The use of potential nonrespondents for studying nonresponse bias. Pacific Sociological Review, 13, 103-109.

Filion, F. (1975-76). Estimating bias due to nonresponse in mail surveys. Public Opinion Quarterly, 39, 482-492.

Filion, F. (1976). Exploring and correcting for nonresponse bias using follow-ups of nonrespondents. Pacific Sociological Review, 19, 401-408.

Fitzgerald, R., & Fuller, L. (1982). I hear you knocking but you can't come in. Sociological Methods and Research, 11, 3-32.

Fowler, F., Jr. (1984). Survey research methods. Applied Social Research Methods Series, 1. Beverly Hills: Sage.

Goyder, J. (1982). Further evidence on factors affecting response rates to mailed questionnaires. _American Sociological Review_, _47_, 550-553.

Goyder, J. (1985). Face-to-face interviews and mailed questionnaires: The net difference in response rate. _Public Opinion Quarterly_, _49_, 234-252.

Hansen, M., & Hurwitz, W. (1946). The problem of nonresponse in sample surveys. _Journal of the American Statistical Association_, _41_, 517-529.

Hawkins, D. (1975). Estimation of nonresponse bias. _Sociological Methods and Research_, _3_, 461-488.

Heberlein, T., & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. _American Sociological Review_, _43_, 447-462.

Henricks, W. (1956). _The mathematical theory of sampling_. New Brunswick, NJ: Scarecrow Press.

Hilgard, E., & Payne, S. (1944). Those not home: Riddle for pollsters. _Public Opinion Quarterly_, _8_, 254-261.

Hippler, H., & Seidel, K. (1985). Schriftliche befragung bei algemeine bevolkerungsstichproben: Untersuchungen zur Dillman'schen total design method. [Mail surveys of the general population: A test of Dillman's total design method.] _Zumanachrichten_. Mannheim: Zentrum fur Unfragen und Analysen.

Hoinville, G., Jowell, R., Airey C., Brook, L., Courtenay, G., Hedges, B., Kalton, G., Morton-Williams, J., Walker, D., & Wood, D. (1978). _Survey research practice_. London: Heinemann Educational Books.

Homans, G. (1961). _Social behavior: Its elementary forms_. New York: Harcourt, Brace & World.

Houston, M., & Jefferson, R. (1975). The negative effects of personalization on response patterns in mail surveys. _Journal of Marketing Research_, _12_, 114-117.

Jolliffe, F. (1986). _Survey design and analysis_. Ellis Horwood Series Mathematics and its Applications. Chichester, England: Ellis Horwood Limited.

Kanuk, L., & Berenson, C. (1975). Mail surveys and response rates: A literature review. _Journal of Marketing Research_, _12_, 440-453.

Kish, L. (1965). _Survey sampling_. New York: John Wiley & Sons.

RASS BIAS REFERENCES

Kish, L., & Hess, I. (1959). A "replacement" procedure for reducing the bias of nonresponse. American Statistician, 8, (October), 17-19.

Lagay, B. (1969-1970). Assessing bias: A comparison of two methods. Public Opinion Quarterly, 33, 615-618.

Linsky, A. (1975). Stimulating responses to mailed questionnaires: A review. Public Opinion Quarterly, 39, 82-101.

Lowe, F., & McCormick, T. (1955). Some survey sampling biases. Public Opinion Quarterly, 19, 303-315.

Lundberg, G., & Larsen, O. (1949). Characteristics of hard-to-reach individuals in field surveys. Public Opinion Quarterly, 13, 487-494.

Mayer, C. (1964). The interviewer and his environment. Journal of Marketing Research, 1, 24-31.

McCarthy, P. (1980). Development of common sampling and surveying methods. Working Paper No. 2, November. ADM Associates.

McCarthy, P. (1981). Development of common sampling and surveying methods for California utilities. Working Paper No. 10, June. ADM Associates.

McCarthy, P. (no date). Use of kWh stratification for reducing variances in end use forecasting. Prepared for the California Energy Commission by ADM Associates.

Moser, C. (1958). Survey methods in social investigation. London: Heinemann Educational Books.

Moser, C., & Kalton, G. (1972). Survey methods in social investigation (2nd ed.). New York: Basic Books.

Nederhof, A. (1981). Some sources of artifact in social science research: Nonresponse, volunteering, and research experience. Unpublished doctoral dissertation, University of Leiden, The Netherlands.

Ognibene, P. (1971). Correcting nonresponse bias in mail questionnaires. Journal of Marketing Research, 8, 233-235.

O'Neil, M. (1979). Estimating the nonresponse bias due to refusals in telephone surveys. Public Opinion Quarterly, 43, 218-232.

Politz, A., & Simmons, W. (1949). An attempt to get the "not-at-homes" into the sample without callbacks. Journal of the

American Statistical Association, 44, 9-31.

Schuman, H., & Kalton, G. (1985). Survey methods. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology: Vol 1. (pp. 635-698). New York: Random House.

Simmons, W. (1954). A plan to account for "not-at-homes" by combining weighting and call-back. Journal of Marketing, 19, 42-53.

Smith, T. (1983). The hidden 25 percent: An analysis of nonresponse on the 1980 general social survey. Public Opinion Quarterly, 47, 386-404.

Srinath, K. (1971). Multiple sampling in nonresponse problems. Journal of the American Statistical Association, 66, 583-586.

Stephan, F., & McCarthy, P. (1958). Sampling opinion: An analysis of survey procedure. New York: Wiley.

Stephenson, C. (1979). Probability sampling with quotas. Public Opinion Quarterly, 43, 477-496.

Stinchcombe, A., Jones, C., & Sheatsley, P. (1981). Nonresponse bias for attitude questions. Public Opinion Quarterly, 45, 359-375.

Swanson, D. (1986). Missing survey data in end-use energy models: An overlooked problem. The Energy Journal, 7, 149-157.

Thibaut, J., & Kelley, H. (1959). The social psychology of groups. New York: Wiley.

Tull, D., & Albaum, G. (1973). Survey research: A decisional approach. New York: Intext Educational.

U. S. Bureau of the Census. (1967). The current population survey - a report on methodology: Technical paper number 7, 1963. Washington, D.C.: U. S. Government Printing Office, 1967.

Warwick, P. & Lininger, A. (1975). The sample survey: theory and practice. New York: McGraw-Hill.

Wasson, C. (1964). Descriptive research: Observational and survey methods. In The strategy of marketing research (pp. 130-164). New York: Appleton-Century-Crofts.

Wayne, I. (1975-1976). Nonresponse, sample size, and the allocation of resources. Public Opinion Quarterly, 39, 557-562.

Wilcox, J. (1977). The interaction of refusal and not-at-home sources of nonresponse bias. Journal of Marketing Research, 14, 592-597.

Wiseman, F. (1976). A reassessment of the effects of personalization on response patterns in mail surveys. Journal of Marketing Research, 13, 110-111.

Yates, F. (1981). Sampling methods for censuses and surveys (4th ed.). New York: Macmillan.

Yu, J., & Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. Journal of Marketing Research, 20, 36-44.

Version: 2/16/88

Energy Audit Programs and Household Energy Retrofit Activity

Abbreviated Report for CEC/UERG

by

Trudy Ann Cameron and Matthew Wright*

Department of Economics
University of California, Los Angeles

ABSTRACT

Sometimes policy makers seek to influence economic activity by providing information, rather than by manipulating relative prices. We formulate a utility-theoretic model for households' decisions to install attic insulation with and without participation in an energy audit program. A *joint* discrete dependent variable model (with FIML estimation) is employed to correct for selectivity bias in assessing program effects. We find that (i.) self-selection bias gives the audit program roughly twice the credit it deserves, and (ii.) policy measures designed to influence retrofit costs or energy prices may have more-discernible *direct* effects on retrofit activity than do audit programs.

* We acknowledge the helpful comments of James Harrigan, Michael Waldman, and seminar participants at UCLA and at the Center for Energy and Environmental Studies at Princeton University. Research support was provided by the California Energy Commission (Standard Agreement #300-85-008, Program in Analysis of California Energy Consumption).

Energy Audit Programs and Household Energy Retrofit Activity
Abbreviated Report for CEC/UERG

Trudy Ann Cameron and Matthew Wright
Department of Economics, UCLA

1. Introduction

"Energy audit" programs have been widely advocated. By providing reliable information, they are intended to increase households' energy conservation retrofit activity even though the actual costs and benefits are unchanged. How successful have energy audit programs been in stimulating retrofit activity? And how do audit programs compare with alternative market-oriented methods for encouraging greater retrofit activity?

2. Modeling Energy Audits and Household Conservation Retrofits

We use a fully utility-theoretic specification for a household's simple discrete choice regarding whether or not to install attic insulation. A recent labor economics model of program participation (Bjorklund and Moffitt, 1987) can be adapted to our discrete choice case. In the context of energy retrofit decisions, the traditional continuous earnings variable (Y_i) is replaced by an unobservable "propensity to retrofit" variable, which we interpret as the "utility difference due to retrofitting." Assume that each household's retrofit decision is based on the level of indirect utility it expects to enjoy with and without the retrofit. Households derive utility from "space heating," "space cooling," and a composite of "all other goods and services." We view a house as a structural shell which retains heat (during space heating) and prevents heat gain (during cooling). Fuel (or electricity) is required to operate heating or cooling equipment, but fuel inputs do not themselves confer utility. Only when they are combined with the capital services of the heat generating (or removal) equipment and the structural shell do they result in the "space comfort" outputs which do provide utility.

As in Cameron (1985), we first consider each household's expenditures on heating and cooling fuel. Demands for these inputs will be derived from their demands for space heating and space cooling:

(1) heating: $(PF_h)f_h = PF_h\ (1/E_h)\ U\ (\ HDD_i\ -\ HDD_o)$

cooling: $(PF_c)f_c = PF_c\ (1/E_c)\ U\ (\ CDD_o\ -\ CDD_i)$

where PF_h and PF_c are the per-unit prices of the household's primary heating fuel and cooling fuel, respectively; f_h and f_c are the nominal quantities of heating and cooling fuel; E_h and E_c are the approximate efficiencies of the heating and cooling equipment; U is the total "lossiness" of the structural shell (in BTU's per degree day); HDD_o and HDD_i are outside and inside heating degree days per year; CDD_o and CDD_i are outside and inside cooling degree days per year.[1,2] Decisions regarding the two types of capital--heat generating (or removal) equipment and the structural shell--are considered medium-to-long-run investment decisions. The amount of temperature compensation is a short-run consumption decision (i.e. thermostat-setting behavior). Local variations in climate will be the primary determinant of the amount of heating (HDD_i - HDD_o) and cooling (CDD_o-CDD_i) households actually consume.

Indirect utility will depend on the household's opportunity sets under each retrofit alternative, which reflect income net of annualized[3] retrofit costs, Y - C, and the "prices" of marginal units of space heating or space cooling: $p_h = PF_h\ (1/E_h)\ U$ and $p_c = PF_c\ (1/E_c)\ U$.

The total U factor in each marginal price will depend upon the size of the house, its configuration, and the materials from which it is constructed. But it will also depend upon the household's decision regarding whether or not to retrofit for energy conservation. By retrofitting, the household gives up some income but also decreases its prices of heating and cooling by reducing the lossiness of the structural shell. If households perceive that the maximum level of utility they can attain *with* the retrofit exceeds that attainable under the status quo, they retrofit. From each household's decision, we can infer something about the configuration of these households' indirect utility function, V*.

A household will retrofit if $V^*(Y - C,\ p_h^a,\ p_c^a)$ exceeds $V^*(Y,\ p_h^b,\ p_c^b)$, where C is the cost of the retrofit measure and the superscripts on the prices

---

[1] These expenditures are distinct from the total fuel expenditures which would appear on the utility bill, since these fuels are frequently used for cooking or water heating as well.

[2] "Inside" degree days can theoretically be positive or negative, depending upon whether the household over- or under-compensates for the total number of outside degree days by its thermostat-setting behavior. Since this behavior is unobserved, however, we end up setting inside degree days to zero in practice.

[3] We must arbitrarily select a planning horizon for households. Four years seems reasonable, so we take annualized costs to be .25 times total cost. The results are not particularly sensitive to this choice.

denote the "after-retrofit" (^{a}) and "before-retrofit" (^{b}) prices of space heating and space cooling with $p_h^a < p_h^b$ and $p_c^a < p_c^b$. We ignore for now the effect of the audit program. However, we cannot fully observe the determinants of V*, so we assume that utility can be decomposed into an observable component, V, and an unobservable normal error component, ϵ. We can then write:

$$(2) \qquad \Pr(\text{retrofit}) = \Pr [\ V(Y - C,\ p_h^a,\ p_c^a) + \epsilon_1 > V(Y,\ p_h^b,\ p_c^b) + \epsilon_0\ ],$$
$$= \Pr [\ (\epsilon_0 - \epsilon_1) < V(Y - C,\ p_h^a,\ p_c^a) - V(Y,\ p_h^b,\ p_c^b)\ ]$$

where the error difference $(\epsilon_0 - \epsilon_1)$ is now treated as a single normally distributed random variable.

Our full retrofit model with its discrete outcome variable can be made explicit as follows. For each household i, unobservable variable R_i*, equal to $V(Y\text{-}C, p_h^a, p_c^a) - V(Y, p_h^b, p_c^b)$, is the amount by which the household's anticipated indirect utility *with* the retrofit exceeds that without it. The observable manifestation of R_i* is R_i. If $R_i^* > 0$, the household retrofits ($R_i = 1$); otherwise, the household does not retrofit ($R_i = 0$). T_i* is an unobservable continuous variable which we will call the household's "propensity to take an energy audit." T_i takes on a value of one if T_i* is positive (audit) and zero if T_i* is negative (no audit). The exogenous variables affecting the retrofit decision are X_i; the benefits of an audit are Z_i, and the costs of an audit are W_i. The structure of the model is:

$$(3) \qquad R_i = 1 \text{ if } R_i^* > 0; \quad R_i = 0 \text{ otherwise}$$

$$(4) \qquad R_i^* = x_i\beta + \alpha_i T_i + \epsilon_i$$

$$(5) \qquad T_i = 1 \text{ if } T_i^* > 0; \quad T_i = 0 \text{ otherwise}$$

$$(6) \qquad T_i^* = \alpha_i - \phi_i, \text{ where}$$

$$(7) \qquad \alpha_i = Z_i\delta + u_i$$

$$(8) \qquad \phi_i = W_i\eta + v_i$$

Without an audit, households face uncertain costs and savings from retrofitting. Because retrofit savings and cost information is commonly given in per-square-foot terms, total uncertainty is likely to be proportional to the square footage of the house. We assume that the *expected* values of retrofit costs and savings match the "certainty" values provided by the audit but that, because of uncertainty, retrofitting without an audit is a "gamble." With risk aversion, the utility from the gamble will be lower than that from the expected levels with certainty. Thus, we assume that *ex ante* anticipated post-retrofit indirect utility differs systematically between audited and non-audited households by an additive "uncertainty reduction" term, proportional to square feet. The coefficient on this variable should be positive.

But the decision to take an energy audit is not necessarily exogenous. Volunteering to take an audit may imply relatively stronger "energy awareness" which may also be positively correlated with retrofit activity. Self-selection into the audit program could seriously bias the audit coefficient in a simple discrete choice model for the retrofit decision.

Our submodel for audit program participation assumes that there are both benefits and costs associated with taking an audit. The increase in expected post-retrofit indirect utility due to uncertainty reduction is the primary "benefit" of an audit (α_i). The costs (ϕ_i) reflect the inconvenience of *arranging* for the retrofit--someone must be at home to admit the auditors and their equipment. Thus, the opportunity cost of time (i.e. income) will inversely affect the probability of taking an energy audit. It is also costly to *find out* about the advantages of audits. Less-educated households may be less well-informed, so the educational attainment of the household head is also included.

See our longer paper (Cameron and Wright, 1987) for details of the reduced form and of the stochastic structure of this model. These assumptions allow the unknown parameters to be estimated by full information maximum likelihood techniques.

3. Data

We utilize a dataset generously provided by the Los Angeles Department of Water and Power: their *Residential Energy Survey* for 1983. Our estimating sample includes 969 respondents living year-round in resident-owned, single family detached homes.

The "program" participation question on the survey reads: "Have you had a home energy survey of your residence?" The retrofit question asks respondents to indicate whether they added ceiling insulation "last year," "prior to a year ago," "never," or whether they were "not sure." The two positive responses were merged and households responding "not sure" were dropped from the sample.[4]

Table I summarizes the acronyms, means, and standard deviations of the constructed variables used for estimation. For the unweighted sample, 313 households neither took an audit nor retrofitted, 134 did both, 63 took an audit but did not retrofit and 459 retrofitted without taking an audit. The observed relative frequency of retrofitting given that an audit *was* taken was .680, as opposed to a relative frequency of .595 when an audit *was not* taken.[5]

4. <u>A Quadratic Indirect Utility Function with Sociodemographic Shifters</u>

Our basic indirect utility function is quadratic:

$$V = \beta_1 Y + \beta_2\, p_h + \beta_3\, p_c + \beta_4\, Y^2 + \beta_5\, p_h{}^2 + \beta_6\, p_c{}^2 + \beta_7\, Y\, p_h + \beta_8\, Y\, p_c + \beta_9\, p_h\, p_c \tag{15}$$

However, it is the difference between (expected) utility *after* the potential retrofit and (known) utility *without* the retrofit which is presumed to determine retrofit activity:

$$\begin{aligned} \Delta V = V_a - V_b = \beta_0 &+ \beta_1(Y\text{-}C - Y) + \beta_2(p_h{}^a - p_h{}^b) + \beta_3(p_c{}^a - p_c{}^b) \\ &+ \beta_4[(Y\text{-}C)^2 - Y^2] + \beta_5[(p_h{}^a)^2 - (p_h{}^b)^2] + \beta_6[(p_c{}^a)^2 - (p_c{}^b)^2] \\ &+ \beta_7[(Y\text{-}C)p_h{}^a - Y\, p_h{}^b] + \beta_8[(Y\text{-}C)p_c{}^a - Y\, p_c{}^b] \\ &+ \beta_9(p_h{}^a\, p_c{}^a - p_h{}^b\, p_c{}^b). \end{aligned} \tag{16}$$

The same parameters appear in ΔV as in V.

We have allowed for systematic variation in the indirect utility parameters by examining the data to determine which β coefficients in equation (16) seem to be influenced by household-specific variables. Rudimentary specification searches suggest: $\beta_1 = \gamma_1 + \gamma_2\, HSZ_i$, $\beta_2 = \gamma_3 + \gamma_4\, DNEW_i$, and $\beta_4 = \gamma_5 + \gamma_6\, SENS_i$. The number of household members (HSZ) for a given level of total household income will affect the household's decision-making process with respect to the allocation of income. If the household has not yet experienced the full cycle of seasons in this particular house (DNEW = 1),

---

[4] It is conceivable that the attic insulation might predate the audit, but we must assume that such events are infrequent.

[5] We are implicitly assuming that sufficient time has passed between the audit program and the time of this survey for the full effects of the audit program on retrofit activity to be realized.

Table I

Variable Acronyms and Descriptive Statistics
(n = 969; owner-occupied single family dwellings)

Acronym	Description	Mean or Proportion (Weighted)	(std.dev.)
RETRO	household retrofitted with attic insulation	0.5732	
AUDIT	household had an energy audit	0.1946	
Y	household income ($ 000)	36.66	(22.77)
C_b	retrofit (full) cost ($'000)	1.519	(0.6298)
P_{h_a}	price of heating (before, $/HDD)	0.5651	(0.2211)
P_{h_b}	price of heating (after, $/HDD)	0.2812	(0.09613)
P_{c_a}	price of cooling (before, $/CDD)	0.5852	(0.6593)
P_{c}	price of cooling (after, $/CDD)	0.3105	(0.3166)
HSZ	number of people in household	2.805	(1.389)
DNEW	=1 if occupied house for < 1 year	0.02667	
SENS	=1 if children or elderly present	0.4112	
SQFT	square footage of house ('00)	17.14	(6.997)
EDUC	education of household head (years)	14.39	(2.762)

they may not yet fully appreciate the costs of not retrofitting, or may not have had the opportunity to retrofit. SENS captures differences in temperature variation *tolerance* of different types of households. Those with small children or elderly people present (SENS = 1) tend to keep interior temperatures within a narrower range. Households having only non-aged adults and school age children seem more resilient to temperature variation and also have fewer inactive people around the house during the daytime.

5. Results and Interpretation

Despite the multicollinearity among many of the variables, we still achieve statistically significant parameter estimates for a majority of the parameters of the indirect utility function.[6]

Results for the joint model appear in Table II. We are most concerned about the implications of the selectivity correction process for the estimated coefficient on the AUDIT*SQFT variable, which is interpreted as the "program effect." Simple probit models for the retrofit decision alone imply that this program effect is statistically discernible. But whereas the coefficient is 0.009347 in the simple *quadratic* probit model (not reported in detail in this summary), it is only 0.005092 in the corresponding jointly estimated model, and this revised estimate is not even remotely statistically significant. So even admitting a correlation of less than 0.1 between just *one* of the two pairs of error terms, the naive probit model overstates the effect of the audit on retrofit behavior by almost double.

6. Simulations: the Practical Implications of the Fitted Models

Our primary interest concerns the implications of our model for the predicted level of retrofit activity with and without the audit program. Simulations will be based upon our model's within-sample predictions of behavior.

Of the 969 (weighted) households in our sample, 555.38 actually retrofitted. Our model suggests that 551.00 households are predicted to retrofit. In a counterfactual exercise, we can set AUDIT = 0 for all households. Artificially removing the audits results in 544.91 households retrofitting (lower by only 6.09, a 1.1% decrease).

---

[6] Of the 36 unique pairs represented by the nine utility function variables, 20 display correlations in excess of 0.6, 9 have correlations greater than 0.8 and 4 correlations exceed 0.9. Six variables are also strongly correlated with SQFT.

Table II

Parameter Estimates, Weighted Sample
(t-statistics in parentheses)
(n = 969;
single-family, owner-occupied dwellings)

Variable	Jointly Estimated Model
1. RETROFIT DECISION MODEL:	
constant	0.3423 (1.978)
(Y-C - Y)	3.315 (2.991)
HSZ * (Y-C - Y)	-0.2047 (-2.512)
$(p_h^a - p_h^b)$	-3.346 (-2.042)
DNEW * $(p_h^a - p_h^b)$	2.307 (2.047)
$(p_c^a - p_c^b)$	-0.4917 (-0.9908)
$(Y\text{-}C)^2 - (Y)^2$	-0.0006275 (-0.03491)
SENS * $(Y\text{-}C)^2 - (Y)^2$	-0.01868 (-3.147)
$(p_h^a)^2 - (p_h^b)^2$	5.118 (3.070)
$(p_c^a)^2 - (p_c^b)^2$	0.4177 (0.9389)
$[(Y\text{-}C)p_h^a - Y\ p_h^b]$	-0.03193 (-1.227)
$[(Y\text{-}C)p_c^a - Y\ p_c^b]$	0.02060 (2.959)
$(p_h^a\ p_c^a - p_h^b\ p_c^b)$	-1.398 (-1.969)
AUDIT*SQFT	0.005092 (0.3079)
2. AUDIT DECISION MODEL:	
constant	-1.912 (-6.968)
SQFT	0.01459 (1.990)
EDUC	0.05744 (2.847)
Y	-0.001171 (-0.4595)
3. PARAMETERS FOR JOINT MODEL:	
σ_ϵ(no audit)/σ_e(with audit)	0.9870 (4.174)
σ_{ef}(with audit)	0.06276 (0.2688)
$\sigma_{\epsilon f}$(no audit)	0.0
max log L	-1078.23

The discernible effects of energy audit participation on retrofit behavior are very small. Is an audit program necessarily the most effective way to encourage retrofit activity? Table III compares *audit* effects on retrofit decisions with the apparent effects of *other* changes in the consumption environment.

Using the parameters for the jointly estimated model, we first simulate changes in energy prices, retrofit costs, and incomes. A 5% higher relative price of heating fuel increases retrofit activity, implying an "arc elasticity" of retrofit activity with respect to heating fuel prices of 0.61. Arc elasticity with respect to the relative price of electricity (for cooling) is roughly 0.09. Simulated changes in real incomes suggest an elasticity of only 0.12. If retrofit costs are adjusted without altering incomes, however, an arc elasticity of about -0.58 seems to apply.

We can also use simulations to show that the presence of children or elderly persons in some households contributed to roughly 30 of the retrofits which took place. Likewise, if it were not for the 3% of recent movers who have occupied their dwellings for less than a year, retrofit frequency would have been higher by about five cases.

We can simulate the possible effects on retrofit activity of modest changes in *household* sizes. The arc elasticity is roughly 0.14; increasing average household size by 5% results in nearly four more retrofits in our simulations. A 5% decrease in dwelling sizes results in 13.88 more retrofits being done, implying an arc elasticity of about -0.50.

All of these simulation results based on the quadratic indirect utility specification seem intuitively plausible. They are limited *ceteris paribus* predictions, but they do provide us with helpful insights regarding the identity and relative influence of factors which affect household energy retrofit activity.

7. Conclusions and Caveats

We have presented an analysis of energy audits as an example of program evaluation when the program in question merely generates information (which reduces the uncertainty associated with households' post-action utility levels). We have adapted a suitable theoretical model from the literature on manpower training and earnings effects to the discrete outcome case.

The dataset available for the audit/retrofit application presented in this paper offers only very sparse information on the types of variables one would ideally require to generate the effective prices of heating and cooling

Table III

Summary of Simulations

Simulation	Number Retrofitting	Absolute Change	Implied Arc Elasticity*
fitted model (actual data)	551.00	-	-
no audit participation	544.91	-6.09	-
5% increase in heating costs	567.70	16.70	0.61
5% increase in cooling costs	553.65	2.65	0.09
5% decrease in retrofit costs	566.97	15.97	-0.58
5% decrease in real income	547.72	-3.28	0.12
5% increase in real income	554.25	3.25	0.12
absence of children and elderly persons (SENS = 0)	520.36	-30.64	-
no recent movers (DNEW = 0)	556.46	5.46	-
5% decrease in avg. hhld size	547.18	-3.82	0.14
5% increase in avg. hhld size	554.81	3.81	0.14
5% decrease in house sizes	564.88	13.88	-0.50

* Elasticities computed for 10% changes were the same to within 0.01 in all cases.

with and without energy retrofits. Simple probit models suggest tangible audit program effects, but unobserved characteristics of households apparently make them more likely to retrofit *and* to take an audit. Self-selection into the audit program means that the simple probit model gives the program considerably more credit than it deserves for stimulating retrofit activity.

We should express two caveats. First, this study addresses only the *direct* effects of the audit program. We cannot assess the externalities generated by the program. A household may be induced to retrofit merely because of *publicity* regarding the audit program, not because it participates itself. However, if these effects are important, it would seem that an intensive general public information campaign might be just as effective as detailed audits conducted for individual households. Second, these findings pertain to households in the service area for the Los Angeles Department of Water and Power.

We have used our fitted models to simulate other counterfactual states of the world besides the absence of audit programs. If energy conservation retrofit activity is too low, it appears that the manipulation of retrofit costs and heating fuel prices might be effective ways to encourage a more socially desirable level of retrofits. Household-specific information dissemination programs, such as the energy audits considered here, appear to have a less discernible effect on behavior.

REFERENCES

Bjorklund, Anders, and Robert Moffitt, "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models," *Review of Economics and Statistics*, 69(1), 42-49.

Cameron, Trudy Ann (1985) "A Nested Logit Model of Energy Conservation Activity by Owners of Existing Single Family Dwellings," *Review of Economics and Statistics*, 67(2), 205-211.

Cameron, Trudy Ann and Matthew Wright (1987) "Evaluating Information Programs when Outcomes are Discrete: Energy Audit Programs and Household Energy Retrofit Activity," mimeo, Department of Economics, University of California at Los Angeles.

TRENDS IN UNIT ENERGY CONSUMPTION:

EVIDENCE FROM SDG&E's MIRACLES 4-6

C.W.J. Granger

Chung-Ming Kuan

Matthew Mattson

Halbert White

Department of Economics
University of California, San Diego

1. INTRODUCTION

The goal of the research is to investigate trends in unit energy consumption using data contained in the Miracle 4-6 datasets collected by SDG&E. If such trends can be clearly identified, then they can be used to improve the forecasting performance of the CEC's residential demand model.

Our analysis does not reveal any clear trends in unit energy consumption. This should not necessarily be interpreted as meaning that there are no such trends. It may be that the time period investigated is too short to reveal such trends. It is also apparent in hindsight that the Miracle datasets may not be well suited to the search for such trends.

Instead, our analysis very clearly shows that end use variables have an extremely small explanatory contribution to make for energy consumption models which include lagged consumption as an explanatory variable. Such models arise naturally in the investigation of the trends we are interested in. Our finding suggests that forecasting models which take this fact into account may usefully complement and supplement existing end-use forecasting approaches.

This brief summary is organized as follows. In Section 2 we discuss theoretical considerations leading to the econometric models used in our analysis. In section 3 we describe data characteristics. An overview of empirical results is discussed in Section 4. Section 5 contains concluding remarks.

2. THEORETICAL CONSIDERATIONS

In the most general framework, the energy usage of an appliance (e.g., KWH per year can be considered as a product of appliance efficiency measure (KWH/usage hour), conditional demand for usage hours per year and an appliance ownership dummy, with these three variables determined endogenously. This immediately leads to a complex simultaneous system including a model of energy efficiency measures, a model of conditional demands, and a model of discrete choice for the appliance portfolio.

Because the objective of our project is to explore possible trends in unit energy consumption for a general appliance portfolio, our model must be able to handle a complete appliance stock. However, consideration of a complete appliance portfolio yields a complicated nested logit model for the discrete choice problem which is not empirically tractable (cf. Dubin (1985) Chap. 3) and McFadden (1978)). We content ourselves with treating appliance ownership as exogenous variables, as in Parti and Parti (1980).

The model formulated in our analysis is a specification of conditional demand for energy as a function of relevant socioeconomic and demographic variables. This specification differs from that of Parti and Parti in several respects. First, we treat efficiency measures as unknown parameters which can be estimated. Second, our model explicitly takes into account lagged consumption and some potentially useful demographic variables not considered by Parti and Parti. Finally, we allow for cross-effects of appliance usage by including interactions of appliance ownership variables as explanatory variables. The model is implemented by using yearly rather than monthly observations. Yearly regressions allow us to focus on the short-run effects of the appliance stock on yearly energy usage so that changing energy consumption behavior may be detected by examining the changes of parameters in regression results.

The inclusion of lagged consumption in our model is of great importance in our analysis. Unlike previous work on unit energy consumption which does not consider possible interperiod movement in energy consumption, our model explicitly allows for dynamic structure in the determination of energy consumption. Our model suggests that the regression coefficient of lagged consumption should equal one, an assumption which can be tested empirically.

3. DATA CHARACTERISTICS:

[A] Engineering Data:

The engineering data on unit energy consumption for certain appliances for the years 1980, 1981, 1983, 1986 and 1987 are available from brochures published by SDG&E. These numbers provide useful benchmarks and are summarized in Tables 1 and 2.

Examination of Table 1 shows some weak downward trends for gas appliances, as well as a good deal of random fluctuation, plausibly attributable to noncomparabilities from one study to the next. An exception is for water heaters, which decline from 1.22 therms per day in 1981 to .6 therms per day in 1983 to 0.32 therms per day in 1987. In contrast, gas main heating increases from .77 therms per hour to 1.09 therms per hour. Table 2 also reveals few clear trends for electric appliances. For example, electric water heater usage exhibits an apparent 1/3 reduction from 1980 to 1981 then fluctuates with some increases over the period.

As there are no definite trends in unit energy consumption apparent in the engineering data, we proceed to an econometric analysis, as discussed in the previous section.

The data used in our regression analysis were obtained from SDG&E. The data comprise part of the "Miracle" Series collected by SDG&E for the purpose of obtaining information about household appliance saturations and customer demographics for the area served by SDG&E. We specifically use Miracles 4,5, and 6 for our analysis.

Each Miracle dataset contains customer responses to a questionnaire and the customer's consumption and billing history for natural gas and electricity. The Miracle 4 survey was conducted in 1979 and resulted in 12,380 usable customer questionnaires. The Miracle 5 survey was sent out in 1981 and yielded 8022 usable questionnaires. The Miracle 6 survey was conducted in 1983 and resulted in 7600 usable questionnaires.

Weather data were also obtained from SDG&E, covering ten weather stations around San Diego County. The data consist of daily highs and lows for each of these weather stations. Each customer was assigned a weather station based on their zipcode.

For each data set the data were transformed from raw survey form into usable "dummy variables" for the questionnaire items. Detailed information on specific variable transformations is given in the complete report in separate Appendices. After dropping all observations for which data were missing or unattributable, we arrived at samples of approximate size 8700, 3600 and 4000 for Miracles 4, 5, and 6 respectively in the electrical appliance regressions, and of approximate size 7400, 3100, and 2900 for Miracles 4, 5, and 6 respectively in the gas appliance regressions.

4. OVERVIEW OF RESULTS

We first observe that there is no simple relationship between the mean energy consumption and the weather means. For example, the lowest gas consumption in 1984 (which might be thought to be largely for heating purposes) corresponds to the coldest average winter. Similarly, the warmest summer, again 1984, does not correspond to an exceptional electricity consumption. Thus, there does not exist a strong relationship between annual energy consumption and weather patterns. Although we do not ignore possible weather effects in our analysis, they do not appear to be playing a leading one.

The dependent variable in our regression equations is annual energy consumption in therms or KWH. Explanatory variables used in our model include log income, lagged energy price, and residence square footage. These variables multiply appliance ownership dummies, and they appear in the regressions as interaction terms.

The models used for analysis are as follows.

(i) Model A1 uses only three demographic variables.

(ii) Model A2 uses just the appliances dummies.

(iii) Model A3 uses three demographic variables and appliance dummies.

(iv) Model A4 is the "complete" regression, using demographic variables, appliance dummies and appropriate interactions. A4a uses price interactions, and A4b uses log income interactions. In order to investigate possible trends, we specify lagged consumption as a dependent variable.

(v) Model B0 uses just the lagged consumption and a constant.

(vi) Model B1-B4 correspond to the A-models but with the previous year's consumption also used as an explanatory variable. All models are estimated using an ordinary least square estimation procedure. The standard errors and t-statistics are the White consistent estimates (White (1980)). These estimates are consistent in the presence of heteroskedasticity of unknown form.

As there are many regression results, we only examine the overall picture. Tables 3 and 4 (Tables 8 and 9 in the complete report) give the R^2 values and the standard errors for electric and gas models described above. We observe that:

(i) R^2 values are always higher for electricity regressions than for the corresponding gas regressions.

(ii) R^2 values are consistently larger for the B models (the model with lagged consumption) than for the corresponding A models.

(iii) The simplest B0 model has a higher R^2 value than A4, the most complicated A model.

(iv) Within the B models, there is very little difference in R^2 values.

Some of these observations can be explained by the fact that last year's consumption contains much more explanatory information about this year's consumption than all of the other explanatory variables used in the study. This is the clearest and most relevant result coming from the project. Note that the standard error for the B0 model, which uses only lagged consumption is typically only 1% or 2% different from the best B model, and less than half that of the best A models.

This result is potentially useful for forecasting, as the only explanatory variable in B0 is lagged consumption. The other models contain explanatory variables that will themselves have to be forecast in order to form a forecast of the dependent variable. This is particularly true for the A models. To provide good forecasts, the coefficient on the lagged consumption should be stable. We summarize these coefficients in Table 5 (Table 10 in the complete report).

Examining Table 5 we see that down any column, that is different models for the same fuel and year, the coefficients are very stable. Across years for any single model, the coefficients are relatively stable for electricity consumption but less stable for gas. Using the coefficients for the simplest model, B0, for example, the electricity coefficients have an average of 0.925 and the ratio of maximum/minimum is 1.102; the gas coefficients have an average of 0.883 and max/min is 1.297. There is no evidence of any systematic movement or trend in these coefficients.

It may be no accident that the "instabilities" evident in the gas consumption coefficients coincide with the movement from one survey to the next. Particularly disturbing is the large change evident between 82½ and 83, a period which contains a 6 month overlap and has rather similar weather patterns for the nonoverlapping period. This may indicate survey incompatibilities which may well invalidate any comparisons one might wish to make across survey years.

To investigate the relative importance of the various explanatory variables considered in the project, a sequence of regressions of varying complexity was run for each year. That is, different sets of explanatory variables are used in a variety of regressions. Although we do not present details in this summary, some conclusions from this analysis can be given: (i) The regional dummies, indicating different weather zones, have little explanatory value, particularly for gas consumption. (ii) R^2 values are always higher for electricity models than for gas models. (iii) Residence square footage, among the demographic variables, has the most explanatory value. (iv) Lagged consumption has much more explanatory power than do all other variables.

It is interesting to note how few of the coefficients are "significant," i.e. coefficients have a t-statistic of two or more in magnitude. Lagged consumption has a very high t-value, but, when it is present, very few other high t-values are observed. In terms of the model we considered, the fact that there are so few significant coefficients on the appliance dummies can be interpreted as evidence of essentially no statistically significant change in unit energy consumption from one year to the next. This is in accord with the engineering data. Moreover, it is plausible that it might take many years of strictly comparable survey data (i.e., a panel) in order to detect any appreciable trends in unit energy consumptions. The Miracle data sets, because of inadvertent but inherent differences across surveys (especially as apparent for gas consumption) may not be equal to the task.

5. CONCLUDING REMARKS:

The original goal of this project was to investigate trends in unit energy consumption using the Miracle 4, 5 and 6 datasets collected by SDG&E. Our statistical analysis has found little evidence of any clear trends. Instead, we discovered that by far the most important explanatory and effective predictive variable is past consumption. It is not surprising that past consumption should play a key role. However, what is surprising is the extent to which past consumption dominates all other predictive variables. Regression standard errors are two or three times as great when past consumption is omitted. When demographic or end-use variables are omitted, regression standard errors are affected by less than 1% in many cases. Independent demographic information and end-use information appear to be not of second order importance, but of third order importance when compared with past consumption.

Unfortunately, this conclusion offers little practical guidance in an environment in which forecasts must be made using an inherently end-use oriented model. For such a model, forecasts of appliance saturations and unit energy consumptions are needed to drive the long-range forecasting exercise. The available evidence reveals no clear trend in unit energy consumptions; the engineering information which we presented in Section 3 turns out to be perhaps the most directly relevant practical information in this regard.

Instead, what we have found suggests that the forecasting exercise (at least in the short to medium run) may benefit from formulation and investigation of an approach which explicitly recognizes that demographic and end use variables are summarized more effectively than anything else by past consumption. In its simplest form, such a model would yield aggregate consumption forecasts by summing over forecasts for each household, where forecasts are driven primarily by past consumption behavior. The key elements of such a forecasting model then become population and demographic forecasts, provided that the time series pattern of household consumption within demographic groups is stable. The evidence we find for instability indicates that further study of time series patterns across households is needed before such an approach could be confidently advocated. At the very least, a true panel dataset would be needed to conduct such a study. The Miracle datasets do not appear to be the best vehicle for such investigations.

TABLE 1

GAS APPLIANCES (therms per use)

Appl. / YEAR	198Ø	1981	1983	1986	1987
Range (per burner)	Ø.26/hr	Ø.1Ø25/hr	Ø.11/hr	Ø.Ø94/hr	Ø.219/hr
Oven	Ø.11/hr	Ø.1Ø25/hr	Ø.1Ø/hr	Ø.Ø94/hr	Ø.Ø77/hr
Dryer	Ø.34/hr	Ø.1538/load	Ø.165/load	Ø.165/load	Ø.118/load
Water Heater	1.22/day	Ø.59/day	Ø.6/day	Ø.39/day	Ø.32/day
Mainheat	Ø.77/hr	1.1Ø2/hr	1.Ø/hr	1.3/hr	1.Ø9/hr
Fireplace	Ø.34/hr	Ø.359/hr	Ø.35/hr	Ø.35/hr	n.a.
Pool Heater	1/hr	3/hr	3/hr	n.a.	n.a.
Spa Heater	n.a.	n.a.	1.75/hr	1.Ø37/ warmup	1.89/ warmup
BBQ Grill	Ø.2/hr	Ø.179/hr	Ø.35/hr	Ø.1886/hr	Ø.1538/hr
Air-conditioner	n.a.	n.a.	n.a.	n.a.	n.a.
Source of Information	Energy Cost Calculator (gold)	Energy Cost Calculator (grey)	Understanding Your Energy Bill/Energy Use Worksheet	Understand-ing What Energy Costs	New Study SDG&E March 16

n.a. = not available

TABLE 2

ELECTRIC APPLIANCES (kilowatts per use)

Appl. / YEAR	1980	1981	1983	1986	1987
Range (per burner)	1.125/hr	1/hr	1.1/hr	1.15/hr	2.637/hr
Oven	1/hr	0.9/hr	0.9/hr	0.973/hr	1.21/hr
Dryer	4.875/HOUR	2.5/load	2.5/load	3.27/load	3/load
Water Heater	12/day	8.5/day	7.6/day	8.485/day	8.62/day
Mainheat	n.a.	n.a.	n.a.	30.6/day	30.6/day
Elect. Blanket	0.75/hr	0.6/hr	0.08/hr	0.088/hr	0.088/hr
Pool Heater	n.a.	n.a.	pump 1.8/hr	n.a.	n.a.
Spa Heater	n.a.	n.a.	pump 1.8/hr	23.7/warmup	28.85/ warmup
Attic Fan	n.a.	n.a.	n.a.	n.a.	n.a.
Air-conditioner	Central 4.63/hr Room 1.75/hr	Central 4.6/hr Room 1.7/hr	Central 4.6/hr Room 1.7/hr	Central 4.955/hr Room 1.86/hr	Central 5.73/hr Room 2.04/hr
Microwave	1.5/hr	1.5/hr	1.5/hr	1.59/hr	1.59/hr

TABLE 2 (continued)

ELECTRIC APPLIANCES (kilowatts per use)

Appl. / YEAR	1980	1981	1983	1986	1987
Dishwasher	1.25/hr	Ø.6/load	Ø.6/load	Ø.62/load	Ø.675/ load
Air Cleaner	n.a.	n.a.	n.a.	Ø.Ø44/hr	Ø.Ø5/hr
Water Bed	5.5/day	5.5/day	5.5/day	4.42/day	4.16/day
Washing Machine	Ø.25/hr	Ø.3/load	Ø.25/load	Ø.265/load	Ø.31/load
Color TV	Ø.375/hr	Ø.3/hr	Ø.33/hr	Ø.35/hr	Ø.2/hr
B&W TV	Ø.25/hr	Ø.3/hr	Ø.25/hr	Ø.265/hr	Ø.Ø5/hr
Refrigerator 17 cubic ft.	4.5/day 16 cubic ft.	5/day	5.1/day	5.575/day	5.6/day
Built in room heater	1.375/hr	1.4/hr	1.4/hr	n.a.	n.a.
Portable Elect. Heater	1.5/hr	1.5/hr	1.5/hr	1.59/hr	1.495/hr
Stereo	Ø.125/hr	Ø.1/hr	Ø.1/hr	Ø.Ø88/hr	Ø.Ø97/hr
Source of Information	Energy Cost Calculator (gold)	Energy Cost Calculator (grey)	Understanding Your Energy Bill/Energy Use Worksheet	Understand-ing What Energy Costs	New Study SDG&E March 16

n.a. = not available

TABLE 3

Electricity Consumption Regressions Summary

R^2 values

Model	80	82	82½	83	84
A1	.3493	.3047	.3058	.3056	.3008
A2	.4192	.4231	.4176	.4154	.4047
A3	.5514	.5248	.5187	.5185	.5063
A4a	.5570	.5829	.5927	.5784	.6363
A4b	.5082	.5067	.5013	.5087	.5001
B0	.921	.928	.934	.880	.897
B1	.925	.929	.934	.885	.898
B2	.923	.931	.935	.884	.900
B3	.926	.932	.936	.888	.900
B4a	.925	.933	.936	.888	.901
B4b	.925	.932	.936	.886	.901

Standard Error of Regression

Model	80	82	82½	83	84
A1	2720	2890	2860	2910	2970
A2	2570	2640	2630	2670	2750
A3	2260	2390	2390	2430	2500
A4a	2250	2250	2200	2280	2150
A4b	2370	2440	2440	2460	2520
B0	949	927	882	1210	1140
B1	925	921	880	1190	1140
B2	934	912	875	1190	1130
B3	918	905	872	1170	1130
B4a	924	903	870	1180	1120
B4b	925	908	874	1180	1130

TABLE 4

Gas Consumption Regressions, Summary

R^2 values

Model	80	82	82½	83	84
A1	.159	.218	.220	.237	.213
A2	.095	.147	.153	.159	.148
A3	.198	.276	.282	.308	.281
A4a	.192	.387	.380	.286	.265
A4b	.190	.271	.273	.296	.273
B0	.873	.851	.867	.828	.858
B1	.874	.853	.868	.836	.858
B2	.875	.853	.868	.832	.859
B3	.875	.855	.869	.839	.860
B4a	.876	.858	.875	.837	.861
B4b	.875	.856	.871	.839	.861

Standard Errors of Regression

Model	80	82	82½	83	84
A1	295	239	233	198	195
A2	306	250	243	208	203
A3	288	231	224	189	187
A4a	289	213	209	192	190
A4b	289	232	226	191	188
B0	114	105	96.2	94.0	82.9
B1	114	104	95.9	91.7	82.9
B2	114	104	96.0	92.9	82.6
B3	114	103	95.8	90.9	82.6
B4a	113	102	93.6	91.8	82.5
B4b	114	103	95.4	91.3	82.3

TABLE 5

Regression Coefficients, Lagged Gas and Electricity Consumption

Electricity

Model	80	82	82½	83	84
	Year				
B0	.861	.942	.929	.942	.949
B1	.824	.923	.917	.904	.931
B2	.839	.918	.911	.896	.916
B3	.810	.899	.899	.864	.905
B4a	.816	.924	.898	.861	.883
B4b	.823	.908	.907	.874	.903

Gas

Model	80	82	82½	83	84
	Year				
B0	.776	.999	.943	.813	.883
B1	.774	.974	.925	.776	.876
B2	.782	.994	.933	.796	.876
B3	.777	.977	.921	.764	.871
B4a	.783	.984	.912	.772	.874
B4b	.780	.979	.924	.769	.873

Social Variation and Residential Energy Use in San Diego:
Exploring Consumption and the CEC Demand Forecasting Model

Loren Lutzenhiser, Bruce Hackett and Howard Schutz
University of California, Davis

Prepared for: UERG/CEC
Program in Analysis of California Energy Consumption
University of California, Berkeley Conference
April 29, 1988

This paper offers an overview of our explorations of the potential for disaggregating the California Energy Commission's (CEC) residential electricity and gas demand forecasting model along social and demographic subpopulation lines. A more complete discussion of our analysis and findings is presented in the draft research report to UERG and Commission staff.

We used the SDGE MIRACLE VI (1983) data set to examine the patterns of energy use of households in San Diego County, where we found evidence of fairly distinct social differences in consumption, controlling for the end-use demand determinates used in the CEC model. Because time and resource constraints forced us to use data from a single year in one locality, it would be premature for us to propose any particular disaggregation scheme along the dimensions that we considered (income, lifecycle, family type, owner/renter status, education, ethnicity, tenure, etc.). Our research was essentially exploratory and was not aimed directly at bringing a new forecasting model to life. Rather, we were interested in determining how adequately the CEC model accounts for the social variation in consumption found in the real world. We found evidence that a good deal of that variation may not be captured in the CEC model --or in the other end-use models with which we are familiar.

Ultimately, the feasibility of disaggregating forecasting models along some socio-demographic lines requires a theory which provides reason to suppose that the sizes and behaviors of subpopulations distinguished on the basis of energy consumption are likely to change, in understandable ways, in the future. On the road to that theory, however, a good deal of empirical work remains. A first step is to determine just how comparable the consumption patterns identified in San Diego are to those which might be found in other parts of California. If understandable social patterns of consumption are found in those comparisons, the next step will involve trying to determine just what the past and future energy use trajectories of those social groups might be.

We encourage those interested in this subject matter to examine the full research report. The present paper only superficially reviews the analysis, and presents with little elaboration two possible social/engineering models of residential gas and electric consumption. Our primary concern here being the research and forecasting implications of those findings.

First a few data quality comments are in order. Comparisons of the MIRACLE VI sample with the 1980 Census indicate that response to the survey was considerably biased toward households with higher incomes and higher levels of education. In addition, persons who had lived at the surveyed addresses for fewer than five years and renters, particularly those living in smaller apartments (one bedroom or less), are significantly underrepresented in the sample. These particular response selectivities, which have been considered in some detail by other CEC/UERG researchers, are also correlated with residential energy consumption and certainly result in overstated estimates of population-level consumption when calculations are made using the unweighted sample.

Forecasters and others who work with these sorts of appliance survey data are also familiar with their high rates of non-response to certain questionnaire items (particularly questions regarding income, dwelling size in square feet and insulation levels). The biases that selective in item responses introduce have also been considered by other CEC/UERG research. Our analysis required that cases missing responses on dwelling size, income, number of household members be excluded from the data set. This adjustment resulted in a reduction of sample size by nearly 50% (from 7600 to less than 4000 cases). The reduction further biased the sample in favor of cases with higher rates of energy consumption, higher incomes and larger dwelling sizes. In addition, when Census neighborhood measures of proportions of minority residents (Asian, Hispanic and Black) and recent immigrants are compared to the neighborhood locations of cases excluded because of item non-response, the evidence suggests that these groups may also be underrepresented in the truncated sample. These missing-data problems are recounted in order to make the point that this subsample, and others like it, are loaded toward a homogeneous subpopulation of residential customers, making it a somewhat unlikely place to find differences of the sort that we were seeking.

Our research strategy consisted of first specifying for each case in our subsample those dwelling, appliance and household characteristics which are used in the CEC residential model --producing a set of 56 end use measures and interaction terms which we identified as the "CEC base model variable set" and which are listed in the appendix to this paper.

We then applied this model to the San Diego data (in essence taking the appropriateness of the CEC specification as given) and estimated its goodness of fit. Because the CEC model forecasts demand for various housing types (single family, multi-family and mobile homes), as well as for four vintage periods[1], we then tested the utility of those housing type/vintage disaggregations in the San Diego case. We also tested other potentially useful architectural and climatic disaggregations; constructed and estimated the relationships of a variety of social, geographic and cultural measures on CEC model-residual consumption; estimated hybrid engineering/social models of electric and gas consumption; examined correlations between social and engineering sets of consumption predictor variables; specified a number of alternative consumption models; and compared the results of aggregation from these models with known population values.

Findings:

- The disaggregation of the CEC model into twelve submodels distinguished by three housing types and four vintages did not provide a substantial improvement in the CEC model's fit to the data. Housing type and vintage variables do make important contributions in the context of a single estimating equation, however.

- The effects of housing vintage are evidenced by the declining average consumption of dwellings built under successively more stringent energy conservation requirements of building codes, suggesting that the CEC model's vintage period specifications are reasonable.

- Significant and large differences were observed between subtypes of multi-family units. Condominiums, townhouses, multi-plexes and apartments differ considerably from one another in average consumption, with some behaving more like single family units. The CEC model might usefully improve its distinction between housing types (currently a simple division of units into "single family detached" and "multi-family" types) by taking these differences into account.

- Use of the CEC model's "miscellaneous end use" formula to calculate estimates of residential lighting and small appliance loads for the MIRACLE VI cases introduced distortions in the data. Empirical work could usefully be done to better estimate "miscellaneous" use, and the effects of price and income on that consumption. It might be plausibly argued that these end uses are relatively price inelastic, although ironically they alone are treated in the CEC model as price or income dependent.

---

[1]The CEC model also differentiates between sub-state climate zones. SDGE's service territory is treated in the model as a single climate zone, with some subsequent gross adjustments.

- The CEC model's specification of the effects of *dwelling size* (square feet) on heating and cooling loads, and of *household size* (number of persons) on water heating, cooking, washing, drying, dish washing and related loads, do not exhaustively account for the correlation of these variables with either electricity or natural gas consumption.

- The effects of household size on electricity and gas consumption are non-linear in the sample and are unlikely to be linear in the population.

- Specification of the effect of household size on consumption is improved by the addition of information about ages of household members, suggesting that household size measure is also related to lifecycle stage and family form. Quite different consumption patterns were noted between household types, with the CEC model specification significantly overestimating the consumption of some groups and underestimating that of others.

- A number of social variables were found to be significant predictors of electricity and gas consumption, *after controlling for the engineering specifications of the CEC model.*[2] Household composition, income, education levels, owner/renter occupancy status, length of residence and geographic location were all found to be significant predictors of variations in consumption which were not captured in the CEC model.

- Some cultural and social class dimensions associated with census tract-level measures of ethnicity, immigration and migration, workforce status, occupation and nativity variables were found to be related to gas and electricity consumption.

- A hybrid social and engineering model may be indicated. Aggregate performance of the present CEC model represents a considerable improvement, however, over aggregation from *mean* consumption values. Use of the CEC sampling strata response weights for the 1983 SDGE survey did not measurably improve aggregate estimation.

Social/Engineering Models:

We turn briefly now to the socially-amended CEC model, fit to the San Diego data. These gas and electric models necessarily include a large number of terms, and therefore unfortunately defy simple interpretation. The models, along with a variable name dictionary and an inventory of the CEC model component terms, are included in the appendix to this report.

[2]When social terms were entered in equations which already contained all of the CEC model variables.

We will not discss the coefficients for the appliance end use estimates (the UECs) here, limiting our comments to the social and demographic variables. We would note, however, that while the UEC estimates in these models are adjusted for behavioral/social variations between households, those values are not substantially different from those produced by the CEC model variable set alone, nor do they differ greatly from the estimates that the CEC model actually uses.[3] The intercept term in the electric equation (1022 kWh), which represent non-frost free refrigerator consumption, as well as miscellaneous consumption not accounted for by the other predictors, does not differ dramatically from the CEC model's "miscellaneous end use" estimate minima of 700 kWh and 1200 kWh, for multi- and single family dwellings.

The variation in consumption accounted for by the social variables is fairly large. Although included throughout the CEC model variable set, *dwelling size* again added to the model is strongly associated with electricity consumption, accounting for an additional 952 kWh/year/1000 square feet of dwelling floor area. Controlling for the CEC model's specification of household size (in interactions with cooking, washing and hot water end uses), *household lifecycle groups* differ dramatically from one another in their annual consumption. Couples, singles, small young families and families with relatively large numbers of children consume less than the middle-aged/middle-sized nuclear families who form the reference category. Multiple-adult groups, families with adults over 65 years of age, and families with adult children consume more than the reference group.

Income is also strongly associated with consumption, with the lowest income groups using less, and the higher income groups considerably more electricity than the middle income ($20-30,000/year) reference group. The relatively small differences between most income categories, with much higher consumption levels estimated for higher income households, suggests that the CEC model's engineering variables better fit the middle classes than the economic extremes. This pattern appears to involve conservative energy use behavior at lower income levels and very high levels of energy consumption (produced perhaps both by the presence of unmeasured appliances and by the "non-conservative" use of energy) among higher income households. It may better support a social class interpretation of the effects of income on consumption than it does a continuous and gradual income elasticity of demand interpretation. Our efforts to further specify the sources of this very high consumption at high income levels included estimating models with terms interacting those income levels with the presence of air conditioning, and with air conditioning in inland mini-climate zones. Those attempts did not, however, support a straightforward "high income = high air conditioner use" explanation.

---

[3]For example the CEC model estimate a color television UEC of 627 kWh, while the model variable set estimates it at 750 kWh and the socially-amended model sets it at 534 kWh.

In other social measures, we found that on average renters consumed somewhat more electricity than did homeowners (perhaps representing the lower overall quality and energy efficiency of rental housing). Education is also moderately related to consumption, with lower education levels associated with higher consumption. This finding might be interpreted as the effect of a correlation of lower education with poorer housing quality, as readily as of evidence of some sort of an "more education = more conservation behavior" or "more education = better energy information" effect.

The effects of the ethnic proxies are interesting, particularly because the relationships between the groups on consumption are identical for both electricity and gas. For these variables, consumption increases with tract-level proportion of Black population, while it decreases slightly with Hispanic population increase and drops quite strongly as Asian population increases. While these variables were imported from the 1980 Census and are not claimed to represent accurate case-level ethnic identification, their correlation with residential consumption lends some support to the findings of other studies which have used household ethnic/cultural information.[4]

The appendix also includes a table which presents the socially-amended *gas* equation. Once again the UECs fairly closely resemble, with a few exceptions[5], those produced by the CEC model variable set and those used in the simulation model itself. Even after controlling for the CEC model's specification of the effects of household size on gas consumption, variations in household type/size were found to be significantly associated with consumption level. Young singles used less gas than older singles, while the other household types all used more than the reference group. These lifecycle effects are not nearly as pronounced as they are in the case of electricity use. The strong income effect on gas consumption, is probably surprising from an engineering standpoint. Once again the highest income group was found to be consuming at a very high level. Shorter-term residents used less gas, while lower education levels were associated with higher consumption. Finally, the electric pattern of higher tract-level Black populations associated with somewhat higher consumption, with lower levels found in areas of higher Hispanic and Asian concentration, was repeated for gas consumption.

<u>Energy use and social theory:</u>

Why should we be interested in this sort of variation? There are several reasons aside from the simple intellectual desire to understand it. If segments of the residential consumer

[4]We have in mind both our own cross-cultural comparisons of electricity and gas consumption in California apartments and work by PG&E staff on the ethnicity variables included in the 1986 RASS survey.

[5]For example, the estimated gas-heated pool consumption of 165 and 147 therms/year may represent a much more conservative use of gas for pool heating than the CEC model's 878 therm UEC suggests.

population have quite different energy consumption profiles, which may result in quite different demand probabilities, forecasting models which do not take those differences into account at the very least risk misunderstanding and misspecifying demand. Now the misestimation of demand made by methods which, for example, attribute some average heating fuel consumption rate per degree day to a particular sort of heating system without regard for who is using it, might turn out to be relatively minor. We simply have no way of knowing at this juncture.

Our theoretical approach to residential energy consumption does not assume that this social variation is an artifact of the sample, that its effects on UEC estimates are random or that the contours of these groups are transitory, however. It seems reasonable to suppose that not only do households differ from one another in dwelling characteristics, technology, family form and behavior (all factors which influence rates of residential energy consumption) *as members of social groups,* but that these differences are critical to those groups' sustained self-definition. In important ways, the social groups that interest us (defined on the basis of class, kin, culture and lifestyle) *are* the consuming units in society. While most everyone in the U.S. will tell you that they are "middle class", in sociological terms that status is clearly not held by a majority of those who aspire to it. The population is actually composed of a variety of distinguishable, if not always clearly distinct, segments --some of which can be seen to have quite different energy demand characteristics from others. The concept of the demand-side "market segment" closely approximates the sort of conception of consumer subgroups that we have in mind, and which might be considered in possible future disaggregated forecasts. Further research is need, however, to determine just how distinguishable social groups are on the basis of their residential consumption and how predictions about their future consumption habits might best be made and tested.

Prospects for Social Disaggregation:

The sort of forecasting that the CEC and California utilities have undertaken represents a significant public commitment to understanding and anticipating the future shape of energy use in California. Far from simply being "technical" exercises, the sort of studies reported here will grow in policy importance as resource and energy issues come once again to command global attention.[6] The CEC model in its present form seems empirically to do a reasonable job of predicting residential consumption, so that the value of amending or developing alternatives to the model will depend upon a number of factors. The state is heavily invested in the present model and, to date at least, arguments for a

[6]In considering the likelihood of a resurgence of interest in energy planning and research, recall Robert Hirsch's recent "Impending United States Energy Crisis" in *Science* Vol. 235 (20 March 1987): 1467-1473.

radical departure from the end use approach have not been persuasive.

While very little is systematically known about how persons routinely use energy in domestic settings, there are a number of trends in energy research which will benefit forecasting in the future. The social study of technology, the diffusion of technical innovations, the social psychology of artifact and tool use, the study of social movements and the dynamics of the evolution, ebb and flow of collective notions, norms, conventions, habits and convictions --all represent areas of study which may seem at this point to be esoteric, but are in their infancy and are likely to supply insights on residential energy use which will have a good deal of policy-relevance.

We are gaining a knowledge of how technologies develop, how they penetrate the household, and how elastic their adoption rates are (among various social groups), how rates of use vary, and how matters like "efficiency" and "conservation" and "consumption" and "status" are in reality constructed by social actors of different sorts. As we develop these understandings, we are better positioned to help demand forecasting gradually move away from proxy and gross assumption, and toward an at least more descriptive model of domestic energy consumption patterns.

But a considerable amount of empirical work remains before the hybrid social-engineering forecasting model might become a reality. Determining just *what the patterns of variation in consumption are* is the first step. Eventually, our characterizations of those patterns must be supported by ethnographic understandings of how varieties of household consumption actually occur, or just what the forms of conduct are that compose them, but for the near-term we have identified several avenues of research which follow directly from the work reported here. They include:

> A more detailed social analysis of the MIRACLE VI data, including examination of appliance saturations and use rates (UEC estimates) for subpopulations. Some of this work is now in progress under CEC/UERG sponsorship. In the case of our present model, the addition of cases from the two more recent MIRACLE surveys (VII and VIII), would permit comparisons of model performance with better samples and more recent time periods.

> The consumption patterns identified in MIRACLE VI should be compared to data from other California utility service areas, from various time periods. We had initially hoped to work with more than one data set, but encountered data access and logistical problems which limited the analysis to the San Diego 1983 data. To the extent that comparable data is available, the socially-amended CEC model should be fit to data from PG&E, SMUD, Southern California Edison, SoCal Gas and LAWP in order to determine the viability of the patterns reported here.

> Consideration should be given to construction of a single multi-year data set, selected using Census-based sampling strata, which approximates the demographics and consumption of substate utility service territories, as a tool for future research. The General Social Survey (GSS) is a very useful ongoing national longitudinal sample of this sort.

> The U.S. Census contains considerable information on population characteristics along social dimensions which are associated with known utility sample bias tendencies. In the estimation of population end-use saturation rates and UECs, experimentation with alternative sample weighting schemes seems indicated.

> Discussions between CEC staff, the utilities and academic researchers should be undertaken to explore the feasibility of increasing the amounts of social information collected in appliance saturation surveys. Useful, and elsewhere routinely collected, information which is generally lacking in those surveys includes: gender and age of head of household, ethnicity, and employment status of adults. Other sorts of new data collection are also imaginable.

> A simple comparison of utility survey questionnaire item coding with the Census reporting categories would insure the greatest comparability between utility data and the 1990 Census. Discussions regarding coordination of data collection around the time of the Census might produce agreements to optimize the comparability of sample and Census data.

> While the results of our work with Census tract-level data can only be claimed to be suggestive, it opens two potentially important new avenues of research. First, the research reported here provides additional confirmation of the ethnic patterns of consumption found in other studies. A careful examination of ethnicity and consumption in the PG&E-RASS 1986 data (which contains case-level ethnicity identifiers) is clearly indicated. The second area of further work involves mapping of sample, census and consumption information. Social location theory suggests that, particularly in the U.S., socio-cultural groups often occupy distinct geographic territories --indicating the potential utility of the "neighborhood" approach that our use census proxies has taken. The comparison of geographical distribution mappings over time would show trends in population growth, energy flows and utility survey sampling patterns --information which might be of considerable use to forecasting.

> The importance of some theoretical and technical questions raised by the research, may be apparent only to forecasters. For an example, our findings may make a contribution to debates on the issue of whether or not *household size* is an important predictor of energy consumption --a question raised by those who argue that *numbers of accounts* are sufficiently precise measures of "consumption units". A similar issue might involve the question of whether or not the relationship between energy consumption and income one that is best captured by the notion of income

"elasticity"?[7] Or is the image of income as a continuum of increasing ability to satisfy one's desires (a seemingly smooth interval measure) actually *imposed* on a reality in which income and energy consumption are related through a series of sharp discontinuities?[8] Does one view hold a different significance for forecasting than the other?

> The attention of statisticians might usefully be turned to the nature of conditional demand UEC estimation, particularly to the data and methods commonly used to estimate UECs in the regression context. In addition, comparisons of submeter data[9] on appliance consumption rates and ranges, with statistically-derived UEC estimates would considerably increase our knowledge of the nature and empirical variations in appliance-related consumption. The eccentricities of both the conditional demand framework and appliance-metering studies, might be brought to light when UEC estimates were calculated for consumption data for households whose actual appliance consumption levels are known.

> The size and complexity of the state-wide CEC demand model itself stands as an obstacle to its evolution. A version of the model which could more readily accommodate the estimation of varying *demand scenarios* (e.g. with varying changes in rates of occupancy, housing density, household sizes and age compositions, conservation effects, technology and fuel shifts, etc.) would be of considerable use both to the Commission's efforts to determine the present model's sensitivity to future changes, as well as to planners and policy-makers who are concerned with the energy implications of different plans and proposals. Mini-computer and mainframe versions of simulation languages have come into common use in the natural sciences since the time that the original CEC model code was produced. We encourage exploration of limited-scope models which might be used to simulate the macro-model's performance in alternative scenarios.

---

[7] A paradigm which assumes that as the amount of income is increased, the means to satisfy individual preferences are better realized and consumption increases more or less continuously as a proportion of income increase.

[8] Discontinuities associated with the social allocation of opportunities, means, housing, habits and status-bound practices. The empirical question raised by this difference in characterization of the income->consumption relationship is: "What does the connection between money and consumption consist of?" or "How does increasing income express itself in increased consumption?"

[9] The sort now routinely collected in peak load and time-of-use studies.

CEC Model End Use Terms and Value Ranges
(using MIRACLE VI coding)

Space Conditioning

CENTRAL ELEC HEAT	(0/1)	CENT ELEC * SQFT	(1-10)
BASEBOARD ELEC HEAT	"	BSBRD ELEC * SQFT	"
HEAT PUMP	"	HEAT PUMP * SQFT	"
CENTRAL GAS HEAT	"	CENT GAS * SQFT	"
GAS SPACE HEAT	"	GAS SPACE HT * SQFT	"
CENTRAL A/C	"	CENT A/C * SQFT	"
WINDOW-WALL A/C	"		
EVAPORATIVE COOLER	"		
CEILING INSULATION	"		
WALL INSULATION	"		
CAULKING-WTHRSTRIP	"		
WINDOW TREATMENTS	"		

Cooking

ELECTRIC RANGE	(0/1)	ELEC RANGE * P/HH	(1-7)
GAS RANGE	"	GAS RANGE * P/HH	"

Clothes Drying

ELECTRIC DRYER	(0/1)	ELEC DRYER * P/HH	(1-7)
GAS DRYER	"	GAS DRYER * P/HH	"

Water Heating

ELEC WATER HEATER	"	ELEC WAT * P/HH	(1-7)
HEAT PUMP WTR HEATER	"	HTPMP WAT * P/HH	"
SOLAR w/ ELEC BACKUP	"	SOL-ELEC * P/HH	"
GAS WATER HEATER	"	GAS WAT * P/HH	"
SOLAR w/ GAS BACKUP	"	SOL-GAS * P/HH	"
WATER HEATER WRAP	"		
LOW-FLOW SHOWER HEAD	"		

ELEC WATER HEATER * DISH WASHER * P/HH	"
ELEC WATER HEATER * CLOTHES WASHER * P/HH	"
HEAT PUMP WATER HEATER * DISH WASHER * P/HH	"
HEAT PUMP WATER HEATER * CLOTHES WASHER * P/HH	"
SOLAR-ELEC WATER HEATER * DISH WASHER * P/HH	"
SOLAR-ELEC WATER HEATER * CLOTHES WASHER * P/HH	"
GAS WATER HEATER * DISH WASHER * P/HH	"
GAS WATER HEATER * CLOTHES WASHER * P/HH	"
SOLAR-GAS WATER HEATER * DISH WASHER * P/HH	"
SOLAR-GAS WATER HEATER * CLOTHES WASHER * P/HH	"

Refrigeration

FROST FREE FRIDGE	(0/1)
(manual defrost = 0)	

"General" End Uses

COLOR TELEVISION	(0/1)
DISH WASHER ELEC	"
CLOTHES WASHER ELEC	"
POOL PUMP	"
GAS POOL HEATER	"
FREEZER	"
WATER BED HEATERS	(1-4)

"Miscellaneous" End Uses

(lighting, small appliances/tools)
PRICE+INCOME/CONSTRAINED

CEC Model Variable Set

GWAL	GAS WALL HEAT
EHP	ELECTRIC HEAT PUMP
EBB	ELECTRIC BASEBOARD HEAT
ECNT	ELECTRIC CENTRAL HEAT
GCNT	GAS CENTRAL HEAT
GWAL_SQF	GAS WALL HEAT * SQFOOT
EHP_SQF	ELECTRIC HEAT PUMP * SQFOOT
EBB_SQF	ELECTRIC BASEBOARD HEAT * SQFOOT
ECNT_SQF	ELECTRIC CENTRAL HEAT * SQFOOT
GCNT_SQF	GAS CENTRAL HEAT * SQFOOT
CINSUL	DWELLING HAS CEILING INSULATION
WINSUL	DWELLING HAS WALL INSULATION
CAULKING	DWELLING HAS WEATHERSTRIPPING-CAULKING
WINDOW	DWELLING HAS WINDOW TREATMENTS
CNAC	CENTRAL A-C
WNAC	WALL-WINDOW A-C
EVAP	EVAPORATIVE COOLER
CNAC_SQF	CENTRAL A-C * SQFOOT
EWT	ELECTRIC WATER HEATING
HPWT	HEAT PUMP WATER HEATING
GWT	GAS WATER HEATING
SEWT	SOLAR-ELECTRIC WATER HEATING
SGWT	SOLAR-GAS WATER HEATING
EWT_PHH	ELECTRIC WATER HEATING * PERSONS PER HOUSEHOLD
HPWT_PHH	HEAT PUMP WATER HEATING * PERSONS PER HOUSEHOLD
GWT_PHH	GAS WATER * PERSONS PER HOUSEHOLD
SEWT_PHH	SOL-ELECTRIC WATER * PERSONS PER HOUSEHOLD
SGWT_PHH	SOL-GAS WATER * PERSONS PER HOUSEHOLD
TANKINSU	WATER HEATER TANK INSUL
LOWFLOW	FLOW RESTRICTED SHOWER
ERNG	ELECTRIC RANGE
GRNG	GAS RANGE
ERNG_PHH	ELECTRIC RANGE * PERSONS PER HOUSEHOLD
GRNG_PHH	GAS RANGE * PERSONS PER HOUSEHOLD
HAVFREZ	HAVE FREEZER
FFFRIDG	FROST-FREE FRIDGE
EDRY	ELECTRIC DRYER
GDRY	GAS DRYER
EDRY_PHH	ELECTRIC DRYER * PERSONS PER HOUSEHOLD
GDRY_PHH	GAS DRYER * PERSONS PER HOUSEHOLD
WTE_D_H	E WATER * DISH * PERSONS PER HOUSEHOLD
WTH_D_H	HEAT PUMP WATER * DISH * PERSONS PER HOUSEHOLD
WTSE_D_H	SOLAR-ELECTRIC WATER HEATING * DISH * PERSONS PER HOUSEHOLD
WTG_D_H	GAS WATER HEATING * DISH * PERSONS PER HOUSEHOLD
WTSG_D_H	SOLAR-GAS WATER HEATING * DISH * PERSONS PER HOUSEHOLD
WTE_C_H	ELECTRIC WATER HEATING * CLTH * PERSONS PER HOUSEHOLD
WTH_C_H	HEAT PUMP WATER HEATING * CLTH * PERSONS PER HOUSEHOLD
WTSE_C_H	SOLAR-ELECTRIC WATER HEATING * CLTH * PERSONS PER HOUSEHOLD
WTG_C_H	GAS WATER HEATING * CLTH * PERSONS PER HOUSEHOLD
WTSG_C_H	SOLAR-GAS WATER HEATING * CLTH * PERSONS PER HOUSEHOLD
COLORTV	1+ COLOR TELEVISION SETS
WATRBEDS	# OF WATERBEDS
DISHMOTR	DISH WASHER (MOTOR EST.)
CLTHMOTR	CLOTHES WASHER (MOTOR EST.)
POOLPUMP	POOL PUMP EST.
GASPOOL	GAS POOL HEATER
CEC79_83	BUILT 1979-83
CEC75_78	BUILT 1975-78
CECPRE70	BUILT PRE-1975
CEC_SFAM	CEC SINGLE FAMILY DETACHED UNIT
CEC_MFAM	CEC MULTI-FAMILY UNIT

Socio-Demographic Variable Set

KIDS	NUMBER OF CHILDREN IN HOUSEHOLD
KIDSCALE	SCALE INCORPORATING AGE-SIZE-NUMBER OF CHILDREN
ADULTS	NUMBER OF ADULTS IN HOUSEHOLD
HHTYPE2	HOUSEHOLD LIFECYCLE TYPOLOGY
ADLTAGES	GROSS ADULT AGES
YOUNG	YOUNGEST ADULTS
OLDER	MIDDLE-AGED ADULTS
ELDER	OLDEST ADULTS
TENURE	LENGTH OF RESIDENCE
NUMINSRT	SQUARE ROOT OF NUMBER IN HOUSEHOLD
NUMINSRD	SQUARE OF NUMBER IN HOUSEHOLD
INCOME4	ANNUAL INCOME CODED TO CATEGORY MID-POINTS ($s)
INCOME5	ANNUAL INCOME CODED TO CATEGORY MIDPOINTS ($1000s)
INCSQU	INCOME SQUARED ($1000s)
INCLMH	INCOME - 3 CATEGORIES
INC.1	INCOME - 10K & LESS
INC.2	INCOME - 10-20K
INC.3	INCOME - 20-30K
INC.4	INCOME - 30-40K
INC.5	INCOME - 40-50K
INC.6	INCOME - 50-75K
INC.7	INCOME - GREATER THAN 75K
OWNDWEL	HOME OWNER
RENTDWEL	RENTER
TENLT1	TENURE LESS THAN 1 YR
TEN1_3	TENURE 1-3 YRS
TEN4_7	TENURE 4-7 YRS
TEN8_11	TENURE 8-11 YRS
TEN12_19	TENURE 12-19 YRS
TENGT20	TENURE GREATER THAN 20 YRS
MALEPAY	MALE PAYS UTILITY BILLS
FEMPAY	FEMALE PAYS UTILITY BILLS
BOTHPAY	BOTH MALE AND FEMALE PAY UTILITY BILLS
ED_LTHS	LESS THAN HIGH SCHOOL EDUCATION
ED_HS	HIGH SCHOOL GRADUATE
ED_TRADE	ATTENDED TRADE-TECHNICAL SCHOOL
ED_SMCOL	COMPLETED SOME COLLEGE
ED_COLGD	COLLEGE GRADUATE
TWNROW	TOWN-ROW HOUSE
DU_TRI	DU-TRI-4 PLEX
SNGFAM	SINGLE FAMILY DETACHED DWELLING
APT	APARTMENT
CONDO	CONDOMINIUM
MARITIME	MARITIME CLIMATE ZONE
COASTAL	COASTAL CLIMATE ZONE
TRANSI	TRANSITIONAL CLIMATE ZONE
INLAND	INLAND CLIMATE ZONE
KSQFT	DWELLING SIZE IN 1000-SQUARE FEET
YANDJ	SINGLES UNDER 35
OANDE	SINGLES OVER 35
ADLTGRP	3+ ADULTS
COUPLE	COUPLES
SMYNGFAM	SMALL YOUNG FAMILIES
OLDERFAM	OLDER FAMILIES
MANYKIDS	3+ CHILDREN
ELDERFAM	OLDEST ADULT-CHILD FAMILIES
TRANSINL	COMBINED TRANSITION AND INLAND CLIMATE ZONES

Census Tract Location Variables

PLT5YRS	PERCENT OF POPULATION LESS THAN 5 YEARS OF AGE
P5T17YR	PERCENT OF POPULATION BETWEEN THE AGES OF 5 AND 17 YEARS
PLT17YRS	PERCENT OF POPULATION LESS THAN 17 YEARS OF AGE
P18T24YR	PERCENT OF POPULATION 18-24 YEARS OF AGE
PLT25YRS	PERCENT LESS THAN 25 YRS
P25T61YR	PERCENT BETWEEN THE AGES OF 25 AND 61 YEARS
PABROAD	CENSUS TRACT PERCENT ABROAD IN 1975
P62GTYRS	PERCENT 62 YEARS OF AGE AND OLDER
PFBORN	PERCENT FOREIGN BORN
PYNG.ENG	PERCENT UNDER 5 YEARS OF AGE WHO DO NOT SPEAK ENGLISH
PABROAD	CENSUS TRACT PERCENT ABROAD IN 1975
PSAMEHS	PERCENT IN SAME HOUSE SINCE 1975
PHSGRAD	PERCENT HIGH SCHOOL GRAD
PCOLLGT	PERCENT COLLEGE GRAD PLUS
PWRFRCFM	PERCENT FEMALE IN WORK FORCE
PWORKERS	PERCENT BLUE COLLAR WORKERS
PHI.OCC	PERCENT WHITE COLLAR WORKERS
PRENTER	PERCENT OF RENTERS
PBLACK	CENSUS TRACT PERCENT BLACK
PASIAN	CENSUS TRACT PERCENT ASIAN
PHISPAN	CENSUS TRACT PERCENT HISPANIC
CTINCOME	ANNUAL INCOME (1980 DOLLARS)
RENT	RENT PER MONTH (1980 DOLLARS)

Electric Model

Dependent Variable **TELEC** TOTAL KWH - SAMPLE YR
R Square .643

----------------- Variables in the Equation -----------------

Variable	B	SE B	Beta	T	Sig T
EHP	-4249.29	811.69	-4249.29	-5.2	.00
EBB	-183.95	316.06	-183.95	-.5	.56
ECNT	-1973.05	663.60	-1973.05	-2.9	.00
EHP_SQF	3.23	.36	3.23	8.9	.00
EBB_SQF	.49	.23	.49	2.1	.03
ECNT_SQF	1.85	.40	1.85	4.5	.00
CINSUL	-28.71	98.49	-28.71	-.2	.77
CINSULEH	266.94	274.25	266.94	.9	.33
WINSUL	77.51	93.65	77.51	.8	.40
WINSULEH	666.62	275.56	666.62	2.4	.01
CAULKING	-66.86	86.58	-66.86	-.7	.44
CAULKEH	643.19	264.10	643.19	2.4	.01
WINDOW	5.12	104.20	5.12	.0	.96
WINDOWEH	432.86	337.36	432.86	1.2	.19
CNAC	-325.45	356.01	-325.45	-.9	.36
WNAC	670.25	128.12	670.25	5.2	.00
EVAP	994.10	434.81	994.10	2.2	.02
CNAC_SQF	.87	.18	.87	4.6	.00
EWT	1520.44	303.75	1520.44	5.0	.00
HPWT	937.93	969.44	937.93	.9	.33
SEWT	495.70	774.87	495.70	.6	.52
WTE_D_H	126.51	118.60	126.51	1.0	.28
WTH_D_H	315.52	457.77	315.52	.6	.49
WTSE_D_H	230.38	414.55	230.38	.5	.57
WTE_C_H	386.95	127.09	386.95	3.0	.00
WTH_C_H	293.14	458.93	293.14	.6	.52
WTSE_C_H	-46.39	441.27	-46.39	-.1	.91
TANKINSU	-38.82	87.56	-38.82	-.4	.65
TANKE	780.96	278.32	780.96	2.8	.00
LOWFLOW	34.35	84.13	34.35	.4	.68
LOWFLOWE	-339.64	282.81	-339.64	-1.2	.22
ERNG	229.84	182.55	229.84	1.2	.20
ERNG_PHH	93.61	57.50	93.61	1.6	.10
FFFRIDG	518.18	116.65	518.18	4.4	.00
HAVFREZ	1008.73	90.09	1008.73	11.1	.00
EDRY	190.27	205.80	190.27	.9	.35
EDRY_PHH	132.34	63.50	132.34	2.0	.03
COLORTV	533.59	151.48	533.59	3.5	.00
WATRBEDS	569.94	76.18	569.94	7.4	.00
DISHMOTR	117.51	33.98	117.51	3.4	.00
CLTHMOTR	131.52	46.95	131.52	2.8	.00
POOLPUMP	2291.82	131.46	2291.82	17.4	.00
EBBMIS	196.86	225.37	196.86	.8	.38
ECNTMIS	182.14	250.20	182.14	.7	.46
EWTMIS	-36.02	203.48	-36.02	-.1	.85
EWDHMIS	80.85	221.02	80.85	.3	.71
EWCHMIS	-92.76	301.96	-92.76	-.3	.75
FFFMIS	731.12	196.61	731.12	3.7	.00
EDRYMIS	-122.98	126.46	-122.98	-.9	.33
POOLMIS	614.30	253.53	614.30	2.4	.01
CEC75_78	-52.56	119.36	-52.56	-.4	.65
CEC79_83	-501.20	150.43	-501.20	-3.3	.00
CEC_MFAM	-226.32	118.89	-226.32	-1.9	.05
KSQFT	952.12	92.88	952.12	10.2	.00
YANDJ	-771.01	252.87	-771.01	-3.0	.00
OANDE	-793.59	230.28	-793.59	-3.4	.00
COUPLE	-287.91	183.54	-287.91	-1.5	.11
ADLTGRP	147.02	188.49	147.02	.7	.43
SMYNGFAM	-487.62	177.59	-487.62	-2.7	.0C
OLDERFAM	96.62	226.12	96.62	.4	.66
MANYKIDS	-776.26	246.80	-776.26	-3.1	.00
ELDERFAM	463.36	195.40	463.36	2.3	.01
INC.1	-379.97	156.46	-379.97	-2.4	.01
INC.2	-170.26	113.58	-170.26	-1.4	.13
INC.4	-93.03	113.59	-93.03	-.8	.41
INC.5	307.72	139.38	307.72	2.2	.02
INC.6	654.58	147.36	654.58	4.4	.00
INC.7	1659.22	193.62	1659.22	8.5	.00
RENTDWEL	277.22	129.93	277.22	2.1	.03
TENLT1	-281.33	202.31	-281.33	-1.3	.16
TEN1_3	-111.16	109.53	-111.16	-1.0	.31
TEN8_11	284.40	124.03	284.40	2.2	.02
TEN12_19	439.43	131.64	439.43	3.3	.00
TENGT20	321.52	137.41	321.52	2.3	.01
ED_LTHS	99.26	171.42	99.26	.5	.56
ED_HS	323.17	118.83	323.17	2.7	.00
ED_TRADE	155.38	187.22	155.38	.8	.40
ED_SMCOL	223.34	91.06	223.34	2.4	.01
MARITIME	60.80	87.22	60.80	.6	.48
TRANSINL	250.07	119.02	250.07	2.1	.03
PHISPAN	-75.36	523.18	-75.36	-.1	.88
PASIAN	-773.70	1281.51	-773.70	-.6	.54
PBLACK	1178.84	676.10	1178.84	1.7	.08
ETHNOMIS	121.59	104.63	121.59	1.1	.24
(Constant)	1022.85	348.79	1022.85	2.9	.00

Natural Gas Model

Dependent Variable **TGAS** TOTAL THERMS - SAMPLE YR
R Square .435

---------------- Variables in the Equation ------------------

Variable	B	SE B	Beta	T	Sig T
GWAL	23.03	29.45	23.03	.7	.43
GCNT	-20.88	30.34	-20.88	-.6	.49
GWAL_SQF	.03	.01	.03	1.5	.11
GCNT_SQF	.07	.01	.07	4.5	.00
CINSUL	38.53	21.64	38.53	1.7	.07
WINSUL	-10.49	22.22	-10.49	-.4	.63
CAULKING	-6.45	21.62	-6.45	-.2	.76
WINDOW	4.45	27.50	4.45	.1	.87
CINSULGH	-74.94	23.17	-74.94	-3.2	.00
WINSULGH	2.62	23.83	2.62	.1	.91
CAULKGH	-.17	22.93	-.17	-.0	.99
WINDOWGH	-14.52	29.23	-14.52	-.4	.61
GWT	68.49	20.46	68.49	3.3	.00
SGWT	92.83	65.33	92.83	1.4	.15
GWT_PHH	8.45	8.20	8.45	1.0	.30
SGWT_PHH	-105.16	48.63	-105.16	-2.1	.03
WTG_D_H	2.06	3.13	2.06	.6	.51
WTSG_D_H	.75	17.96	.75	.0	.96
WTG_C_H	4.10	5.69	4.10	.7	.47
WTSG_C_H	84.58	45.85	84.58	1.8	.06
TANKINSU	6.06	21.95	6.06	.2	.78
LOWFLOW	-26.86	20.35	-26.86	-1.3	.18
TANKG	-32.91	23.24	-32.91	-1.4	.15
LOWFLOWG	9.52	21.69	9.52	.4	.66
GRNG	-26.69	17.24	-26.69	-1.5	.12
GRNG_PHH	18.79	5.35	18.79	3.5	.00
GDRY	9.66	18.74	9.66	.5	.60
GDRY_PHH	11.21	5.67	11.21	1.9	.04
GASPOOL	147.52	19.57	147.52	7.5	.00
GWALMIS	7.32	39.01	7.32	.1	.85
GCNTMIS	42.76	38.86	42.76	1.1	.27
GWDHMIS	-56.13	30.33	-56.13	-1.8	.06
GWCHMIS	105.35	32.32	105.35	3.2	.00
GDRYMIS	-25.81	11.99	-25.81	-2.1	.03
CEC_MFAM	-65.57	11.36	-65.57	-5.7	.00
CEC75_78	-29.75	12.38	-29.75	-2.4	.01
CEC79_83	-110.54	15.16	-110.54	-7.2	.00
KSQFT	47.62	15.98	47.62	2.9	.00
YANDJ	-46.49	26.22	-46.49	-1.7	.07
OANDE	-20.50	23.14	-20.50	-.8	.37
COUPLE	21.11	17.73	21.11	1.1	.23
ADLTGRP	28.32	17.47	28.32	1.6	.10
SMYNGFAM	17.91	16.69	17.91	1.0	.28
OLDERFAM	15.15	21.32	15.15	.7	.47
MANYKIDS	2.18	22.89	2.18	.0	.92
ELDERFAM	20.08	18.06	20.08	1.1	.26
INC.1	-.66	15.08	-.66	-.0	.96
INC.2	-16.98	11.02	-16.98	-1.5	.12
INC.4	6.21	10.83	6.21	.5	.56
INC.5	32.81	13.14	32.81	2.4	.01
INC.6	51.20	14.21	51.20	3.6	.00
INC.7	147.13	19.71	147.13	7.4	.00
RENTDWEL	4.89	12.90	4.89	.3	.70
TENLT1	-62.15	21.24	-62.15	-2.9	.00
TEN1_3	-13.23	10.82	-13.23	-1.2	.22
TEN8_11	-4.88	12.03	-4.88	-.4	.68
TEN12_19	30.00	12.29	30.00	2.4	.01
TENGT20	17.83	12.65	17.83	1.4	.15
ED_LTHS	19.48	16.07	19.48	1.2	.22
ED_HS	23.33	11.20	23.33	2.0	.03
ED_TRADE	22.60	17.86	22.60	1.2	.20
ED_SMCOL	6.93	8.82	6.93	.7	.43
MARITIME	11.74	8.25	11.74	1.4	.15
TRANSINL	31.73	10.77	31.73	2.9	.00
PHISPAN	-93.74	47.48	-93.74	-1.9	.04
PASIAN	-148.36	114.02	-148.36	-1.3	.19
PBLACK	46.92	58.66	46.92	.8	.42
ETHNOMIS	-4.26	10.14	-4.26	-.4	.67
(Constant)	241.55	37.33	241.55	6.4	.00

Evaluating Biases in Conditional Demand Models Using Validation Data

Project Summary
For UERG-CEC Conference on the
Analysis of California Energy Consumption Data
at Berkeley
April 29, 1988

by
Paul Ong
Nirvikar Singh
Suzanne Holt

Acknowledgements

We are grateful to Michael Gergen and Steven Schleimer for research assistance, and to Jim Mulherin for invaluable advice on computer-related issues. The views expressed herein are our responsibility.

I. Introduction

End-use or conditional demand models have been widely used by both academic researchers and utilities in modelling energy demand. Estimation of the models is typically based on data collected by surveys of users. There are two forms of non-sampling errors in such data: response error and non-response error. The latter leads to the information available being incomplete. This incompleteness is generally not random, but systematic, and this ultimately leads to bias in estimation of the unit energy coefficients when standard regression techniques are used. Response error is akin to the usual problem of measurement error, and leads to bias whether it is systematic or random. An addtional consequence of both types of error is bias in estimating saturation levels.

Non-response error, or sample selection bias, and the resulting bias can be dealt with through a simple two-stage estimation method proposed by Heckman (1979), and applied to energy demand estimation in e.g., Skumatz, Holt, Barnes and Ong (1986) and in HOS (1987). However, there is no such method available for response error, which can only be dealt with by getting better data. Typically this is not possible. However, the PG&E validation survey data overcomes the response error through independent data collection by trained individuals visiting survey households. Hence the validation survey data allows one to check how serious is the bias due to response error. Furthermore the validation survey data overcomes the non-response error. Hence it allows one to check how serious is the bias due to non-response error, and how well the Heckman two-stage method corrects for it.

More precisely, we can estimate three models:

(1) The conditional demand model based on the survey response data estimated by stnadard regression -- the "uncorrected" or conventional model;

(2) The conditional demand model based on the survey response data, estimated by Heckman's two-stage method -- the "corrected" model;

(3) The conditional demand model based on the validation data, estimated by standard regression -- the "true" or valid model.

If there is no misspecification, and the validation data is free of errors, then the estimation of the true model gives unbiased estimates of the demand equation parameters. If the parameter estimates of the corrected model are close to those of the true model, this would suggest that the Heckman method works well and that response error is not a serious problem. If the parameter estimates of the uncorrected model are close to those of the true model, this would suggest that response error and non-response error together are not a serious problem. This might be because neither is, or because the biases tend to cancel each other out. This can happen because the reasons for the biases are different for the two types of error. Hence it also follows that if response error is a serious problem, the corrected model may do worse than the uncorrected model. Only data such as that available through the validation survey can allow one to sort out the different possibilities, and suggest whether the uncorrected or corrected model is better for estimation when validation data is not available.

While models (1) and (2) are the only ones that can be estimated with data as usually collected -- by survey response -- and while model (3) is the "best" model under the assumption of no specification error, two more models can be estimated from the validation data, for the purposes of decomposing the effects of errors in variables and sample selection bias, and for assessing the effect of the Heckman method when there are no errors in variables. These models are:

(4) The conditional demand model based on the validation data -- "informed" subsample only -- estimated by OLS;

(5) The conditional demand model based on the validation data -- "informed" subsample only -- estimated by Heckman's two-stage method.

A comparison of model (4) with the uncorrected model (1), allows one to isolate the effect of response errors in estimating the demand of informed individuals. A comparison of model (4) with the true model, (3), on the other hand, allows one to isolate the effect of sample selection bias in estimating the demand of the entire sample. Also, comparing model (5) with the true model allows one to see how well the Heckman method moves the estimates toward the "true" "best" ones. A further indication of the effects of the Heckman method can be obtained by comparing models (1) and (2) as well. Contrasting this with the "true" comparison -- models (3) and (5) -- gives an idea of how the Heckman method appears to perform in the presence of errors in variables. Including this last comparison, there are hence six pairwise comparisons of the five models. The models and the comparisons are summarized in Figure 1.

Briefly, our results show that response error is a more important source of bias in the conditional demand equations than is non-response error. Furthermore, the Heckman method always move the estimates close to the true model either in the absence or presence of errors in variables. The reason in this case appears to be that it "over-corrects."

On the other hand, the uncorrected model is also not close to the validated model. Whether the uncorrected or corrected model is considered closer to the true one depends on the metric one uses. There seems no single satisfactory measure. However, subjectively looking at the economically more important variables, and assessing the economic significance of differences between estimates, it seems that the Heckman method is a slight improvement.

It may be noted that to the extent that response error is systematic, we may improve on the uncorrected and corrected models by adjusting the data for systematic response biases. This will provide a method suitable for situations where validation data is not available, if one believes that the causes of systematic response biases are stable across data sets. This will not, of course, deal with the bias due to random response error. As an example of systematic response error, those who have ceiling insulation are much more likely to report this correctly than those who do not. Another example is the reporting of the presence of a frost-free upright freezer. Those not having one were much more likely to report this correctly than those who did have it. In these two examples, the probability of a correct response depends on the true "state of the world. " Some prelimanary work suggests that respondent characteristics (e.g. ownership) and residence characteristics (e.g. single family) may also be important.

2. Description of Response and Non-response Errors

In this section we highlight some of the characteristics of response and non-response errors in the validation sample of 901 households. We concentrate on variables that are significant in their impact on energy consumption and perhaps also provide possibilities for conservation, for example the presence or absence of wall and ceiling insulation, square footage and the presence or absence of a blanket on the water heater. Other variables include the age of the dwelling, the number of bedrooms, and the numbers and types of various appliances. The existence of response errors and/or non-response errors for some of these types of variables indicate that people guess answers and that they misunderstand questions. Figure 2 gives a graphic representation of the quality of the suvey data for some of the above-mentioned variables: 'blank' and 'don't know' are combined for convinience. Figures 3 and 4 provide further details for square footage. It may be seen that the patterns in the latter are consistent with perceptual errors and responses tending towards "social norms".

We summarize the responses for both the RASS and validation data for some interesting characteristic variables, and then provide a more detailed breakdown. The means for these variables are provided in table 0. For all the variables, the figure is the proportion in the sample having, or claiming to have the residence characteristic or appliance in question, and the remainder is negative responses.

Table 0.

Variable	RASS	Validation
Ceiling Insulation	.87	.614
Wall Insulation	.69	.60
Water Heater Blanket	.58	.5
Chest Freezer	.15	.16
Upright Freezer	.25	.22
Upright Freezer Frost-Free	.10	.07

Table 1 presents the comparison of the original and validation data for the presence of ceiling insulation. The columns are the original responses, and the rows are the result of the validation survey. Each cell has two numbers. The top number is the frequency of observations in that category. The lower number is the row proportion, hence the relative frequency of each answer given the true state (in Table 1, given either the presence or absence of ceiling insulation).

Table 1

Ceiling Insulation

		Responses			
		Yes	No	Blank/Don't Know	Total
	Yes	504	19	30	553
		0.91	0.03	0.06	1.00
Validation	No	198	83	67	348
		0.57	0.24	0.19	1.00
	Total	702	102	97	901

It is clear from Table 1 that those people who had ceiling insulation were very aware of the fact. On the other hand, those who did not, wrongly believed or said they did more than half the time, and a large proportion professed ignorance. Only a quarter of those without ceiling insulation knew or admitted it.

The results for wall insulation are similar in character, as shown in Table 2. Once again, those who had the insulation and put down a definite answer were very likely to be correct, while those who did not have the insulation were often wrong in their answers. For both states of insulation, the proportion of blank or 'don't know' answers was higher (and by roughly the same amount). This accords well with the fact that the presence of wall insulation is harder to ascertain. What is interesting is the much higher proportion of those wrongly reporting that they had ceiling insulation than those wrongly reporting that they had wall insulation. A plausible explanation is that, while ceiling insulation was not present significantly more often than wall insulation, people more readily assumed that ceiling insulation was present, or thought that it should be there.

Table 2

Wall Insulation

		Responses			
		Yes	No	Blank/Don't Know	Total
Validation	Yes	373 0.69	37 0.07	128 0.24	538 1.00
	No	73 0.20	165 0.46	120 0.34	358 1.00
	Total	446	202	248	896

Turning to another energy-conserving feature, the presence of a water heater blanket (Table 3), we again find similar patterns of response to those above: people with the feature are less likely to be wrong, and less likely to not know, than those without.

Regarding appliances there were, unsurprisingly, small errors in responses concerning variables such as 'type of range or stove', 'presence of a microwave oven', and 'number of refrigerators.' Larger errors were present in answers to questions concerning freezers, for example, people were much more likely to give the incorrect response when the appliance or characteristic was present than when it was absent. This could be a combination of misunderstanding the description of the appliance, and a lack of expectation about its presence (unlike, say, ceiling insulation).

So far, we have concentrated on the relative frequencies of different responses for each true state of the characteristic (e.g. presence or absence of ceiling insulation). In practice, in general we observe responses, but do not have validation survey data. We therefore present below, for illustration, Tables 1a and 2a for ceiling and wall insulation respectively. The frequencies are the same as in Tables 1 and 2, but the proportions are now column proportions, i.e. the proportions correct and wrong of each type of definite respondent, and the proportions of the true state for non-responses.

Table 1a

Ceiling Insulation

		Responses: Yes	No	Blank/Don't Know	Total
	Yes	504 0.72	19 0.19	30 0.31	553
Validation	No	198 0.28	83 0.81	67 0.69	348
	Total	702 1.00	102 1.00	97 1.00	901

We see that the patterns of response are similar in the two cases in that the proportions correct are similar for the different types of insulation. They are also similar for positive and negative responses in each case. However, a non-response in the case of ceiling insulation is much more likely to be a residence without such insulation, while in thecase of wall insulation it is a toss-up between presence and absence of the insulation.

Similar proportions can be calculated and compared for having a water heater blanket. The pattern is quite similar to that of ceiling insulation.

Table 2a

Wall Insulation

		Responses			
		Yes	No	Blank/Don't Know	Total
	Yes	373 0.84	37 0.18	128 0.52	538
Validation	No	73 0.16	165 0.82	120 0.48	358
	Total	446 1.00	202 1.00	248 1.00	896

The descriptions above allow us to state some general conclusions. There are some systematic patterns in response errors and in lack of response. This implies that estimated saturation levels based on survey data may be biased. This is illustrated in Figure 5, where the intermediate estimate, based on the validation data, but with the subsample who responded fully to the survey, allows a decomposition into biases due to sample selection and due to response error. Also, response errors are quite large in some cases. Hence, they are likely to have important effects on estimated energy demand equations. We explore these effects and dealing with them in the next section.

3. Estimation and Comparison of Models

In this section we estimate the following five models and compare them, as outlined in the Introduction:

(1) Conventional or uncorrected model, with RASS data;

(2) Corrected model, with RASS data and Heckman method;

(3) Model with validation data, entire sample;

(4) Model with validation data, informed subsample;

(5) Model with validation data, informed subsample and Heckman method.

3.1 Methodology

All the models are estimated by weighted least squares (WLS), to correct for heteroskedasticity. We include monthly seasonal dummies to deal with serial correlation and seasonality. These dummies also partly deal with the failure of the conventional heating degree day variable to adequately capture the effects of temperature variations on conditional energy demand. The GLS procedure improves the efficiency of coefficient estimates, and removes bias in the computed standard errors, and is desirable for those reasons.

Once the five different models are estimated, the pairwise comparisons discussed in the introduction and diagrammed in Figure 1 can be carried out. Each pairwise comparison is made as follows. Differences in pairs of individual coefficients for each pair of models are calculated. The standard error of each difference is calculated according to the formula

$$SE(diff) = \quad SE(1)^2 + SE(2)^2$$

This formula is an approximation, since it neglects the covariance term, which would be -2cov(1,2) . This is not easily calculable, since the two models are in most cases estimated from different data sets. There are no standard theoretical formulas that we are aware of. However, since the data sets will be positively correlated in general, we would expect the covariance term to be positive. Hence, leaving it out gives an overestimate of the true standard error of the difference of coefficients, and hence an underestimate of the degree of statistical significance of the difference. For comparisons such as models (1) and (2), where both data sets are the same except for a single variable, we would expect the covariance to be high and the standard error to be substantially overestimated. However, in general it is difficult to guess the extent of significance. This is not too much of a problem, however, since ultimately what matters is the comparison of comparisons, i.e. how different models do against the benchmark, "true" or valid model.

3.2 Response and Non-Response Errors

We first examine how the conventional model estimates, which are affected by both response error and non-response error, compare with the estimates based on the validation data. We then decompose the effects of response error and non-response error, by comparing each of the above sets of estimates with those from the validation data restricted to the informed subsample of people. Figures 6 and 7 summarize the comparisons.

(i) Electricity Demand

For the comparison, 24 of the 57 differences in coefficients (exluding seasonal dummies) were statistically significant. These included variables such as those indicating the presence of an upright frost-free freezer, electric heat with a central furnace (multiplied by square footage and heating degree days), electric water heat with a low flow shower head, electric heat for units built after 1978 (multiplied by square footage and heating degree days) and miscellaneous electric appliances (VCR, PC, stereo). 14 of the

24 significant differences (model (1) - model (3)) were negative, indicating no consistent direction of bias. It seems reasonable to say, therefore, that the combined effect of errors in variables and sample selectivity is a substantial bias in the electricity demand equation.

Gas Demand

Comparing the conventional model with RASS data and the model with validation data, in the case of the gas demand equation, excluding the monthly dummies, 11 of 29 differences were significant. These included important variables involving the presence of gas space heat multiplied by square footage and heating degree days, and further multiplied by one of the following variables: whether the unit was built after 1978, whether it had a clock thermostat, and the income of the household. Other significant differences involved variables based on the presence of gas water heat, gas cooking and a gas clothes dryer. 6 of the 11 significant differences were positive: hence there was again no consistent direction of bias. The proportion of significant differences was almost the same as for electricity demand. Again, it seems reasonable to state that the combined effect of errors in variables and sample selectivity is a substantial bias in the gas demand equation.

(ii) Electricity Demand

In this comparison, there were 18 significant differences. 11 of these were also significant in the first comparison, so one can partially attribute those biases in the first comparison to errors in variables. Note that some biases due to errors in variables do not show up when there is sample selection bias as well -- the two effects may indeed mitigate each other.

Gas Demand

In contrast to the electricity demand model, there were as many significant differences as in the first comparison, 11 excluding the seasonal dummies. However, only 5 of these were the same as in the first comparison. Hence, as in the electricity demand case, sometimes biases due to errors in variables are mitigated by biases due to sample selectivity.

(iii) <u>Electricity Demand</u>

In this case, there were 9 significant differences. 4 of these were also significant in the first comparison, while 6 were also significant in the third comparison. Thus in some cases, the combined effect of individual biases shows up in the decomposition in comparisons (iii) and (iv). For some variables, however, there was no evidence of a significant bias when only one cause was present -- however their combination led to a significant bias. The conclusions are in line with theory -- it is hard to predict low factors such as errors in variables and sample selection bias will interact, and the "proof of the pudding is in the eating." In terms of numbers of significant differences, however, is seems that errors in variables bias is a more serious problem than sample selection bias in this case.

<u>Gas Demand</u>

For this comparison, there were only 5 significant differences.. This suggests that sample selection bias is less serious than errors in variables bias for gas demand as well, and that any misspecifica-
tion of seasonal patterns of usage is not due to sample selectivity. 3 of the 5 significant differences were also significant in the first comparison, which looked at combined biases. Of that first comparison, 4 of the significant differences cannot be attributed solely to either errors in variables or sample selectivity: it is their interaction that leads to the bias.

To sum up:

(i) There are important biases due to errors in variables, sample selectivity, and their combined presence.
(ii) In some cases, the combination of errors in variables and sample selectivity reduces bias from that due to a single factor, but in some cases bias is increased.
(iii) On the whole, errors in variables seem a more important source of bias than does sample selectivity.

Finally, a graphic representation of the implication of these conclusions for estimated UES's is in Figure 8.

3.3 Correcting for Sample Selectivity Bias

The results of the previous section suggest that, while response error is a more important source of bias, sample selectivity or non-response error also contributes to bias in the estimates of the parameters of the energy demand equations. Accordingly, we examine here the performance of Heckman's two-step method for correcting for sample selection bias. The results are provisionally discouraging for the use of this method, since the estimates using Heckman's method also appear to be subject to biases. We say provisionally since our conclusions are conditional on assuming there is no misspecification in the model, so that the estimates using the validation data are unbiased. This assumption may not hold, though it is beyond the scope of the present study to examine its validity. Also, the Heckman method does appear to do a little better than the conventional one, when applied to the RASS data.

References

Blumstein, Carl, "Non-Sampling Errors in Surveys of Residential Energy Consumption," UERG paper, University of California, Berkeley, December 1985

Cowing, Thomas G., and Daniel L. McFadden, Microeconomic Modelling and Policy, Orlando, Fl. Academic Press, 1984.

Hay, Joe, Robert Leu, and Paul Rohrer, "Ordinary Least Squares and Sample Selection Models of Health Care Demand: A Monte Carlo Comparison," Stanford University, Hoover Institution, Working Papers in Economics E-85-29, 1985.

Heckman, James J., "Sample Selection Bias as a Specification Error," Econometrica, 47(1): 153-161, January 1979.

Holt, Suzanne, and Paul M. Ong, "RASS Data Evaluation Project," report to Pacific Gas and Electricity Company, March 1986

Holt, Suzanne, Paul M. Ong, and Nirvikar Singh, "UECs and Income Elasticities: Estimates from First and Second Generation End-Use Models", report to the California Energy Commission and the Universitywide Energy Research Group, June 1987

Maddala, G.S., Econometrics, , McGraw Hill, 1977

Madow, William G., Incomplete Data in Sample Surveys, New York: Academic Press, 1983

Parti, Michael, and Cynthia Parti, "The Total and Appliance Specific Conditional Demand for Electricity in the Household", Bell Journal of Economics, 11(1): 309-321, Spring 1980.

Skumatz, L., S. Holt, R. Barnes, and P. Ong, "Truncation Bias in Energy Demand Equations: Effects of Missing Survey Data," in C. Smith, T. Davis, and M. Keneipp, eds., Productivity through Energy Innovation, pp. 525-537, Pergamon, New York, 1986

Swanson, David A., "Missing Survey Data in End-Use Energy Models: An Overlooked Problem," The Energy Journal, 7(3): 149-157, July 1986.

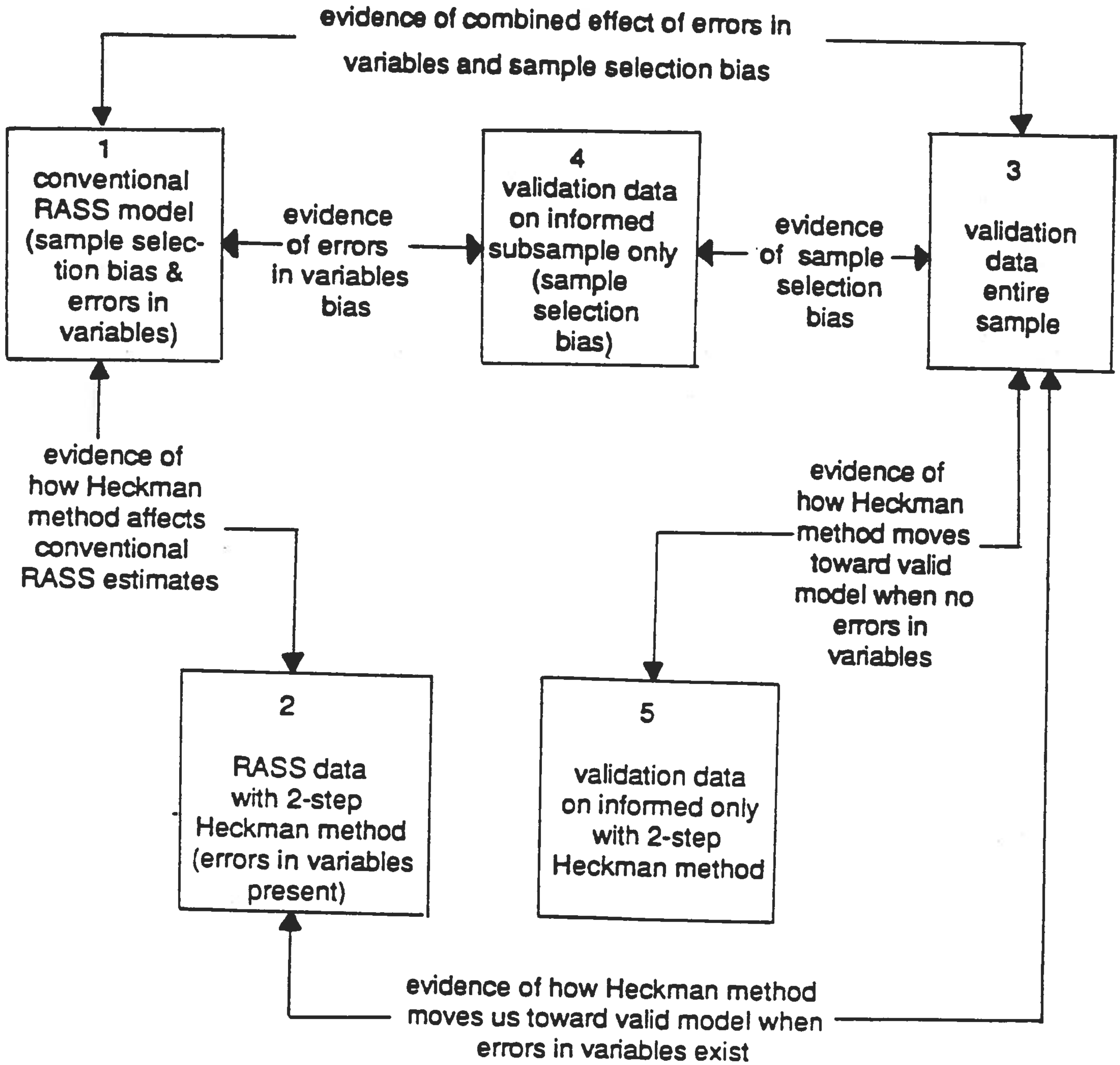

evidence of combined effect of errors in variables and sample selection bias
1
conventional RASS model (sample selection bias & errors in variables)
evidence of errors in variables bias
4
validation data on informed subsample only (sample selection bias)
evidence of sample selection bias
3
validation data entire sample
evidence of how Heckman method affects conventional RASS estimates
evidence of how Heckman method moves toward valid model when no errors in variables
2
RASS data with 2-step Heckman method (errors in variables present)
5
validation data on informed only with 2-step Heckman method
evidence of how Heckman method moves us toward valid model when errors in variables exist

Figure 1

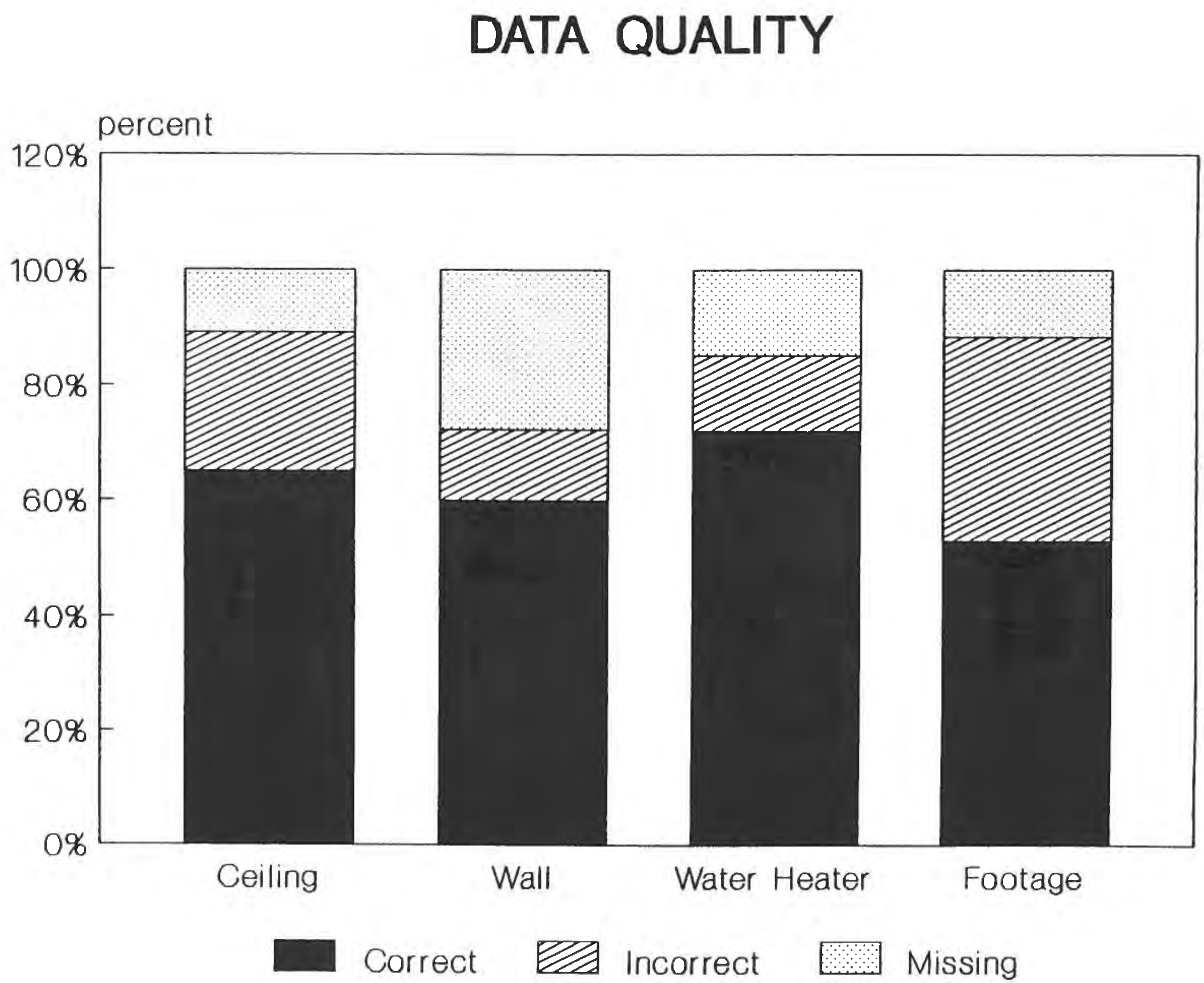

DATA QUALITY
percent
120%
100%
80%
60%
40%
20%
0%
Ceiling
Wall
Water Heater
Footage
Correct
Incorrect
Missing

Figure 2

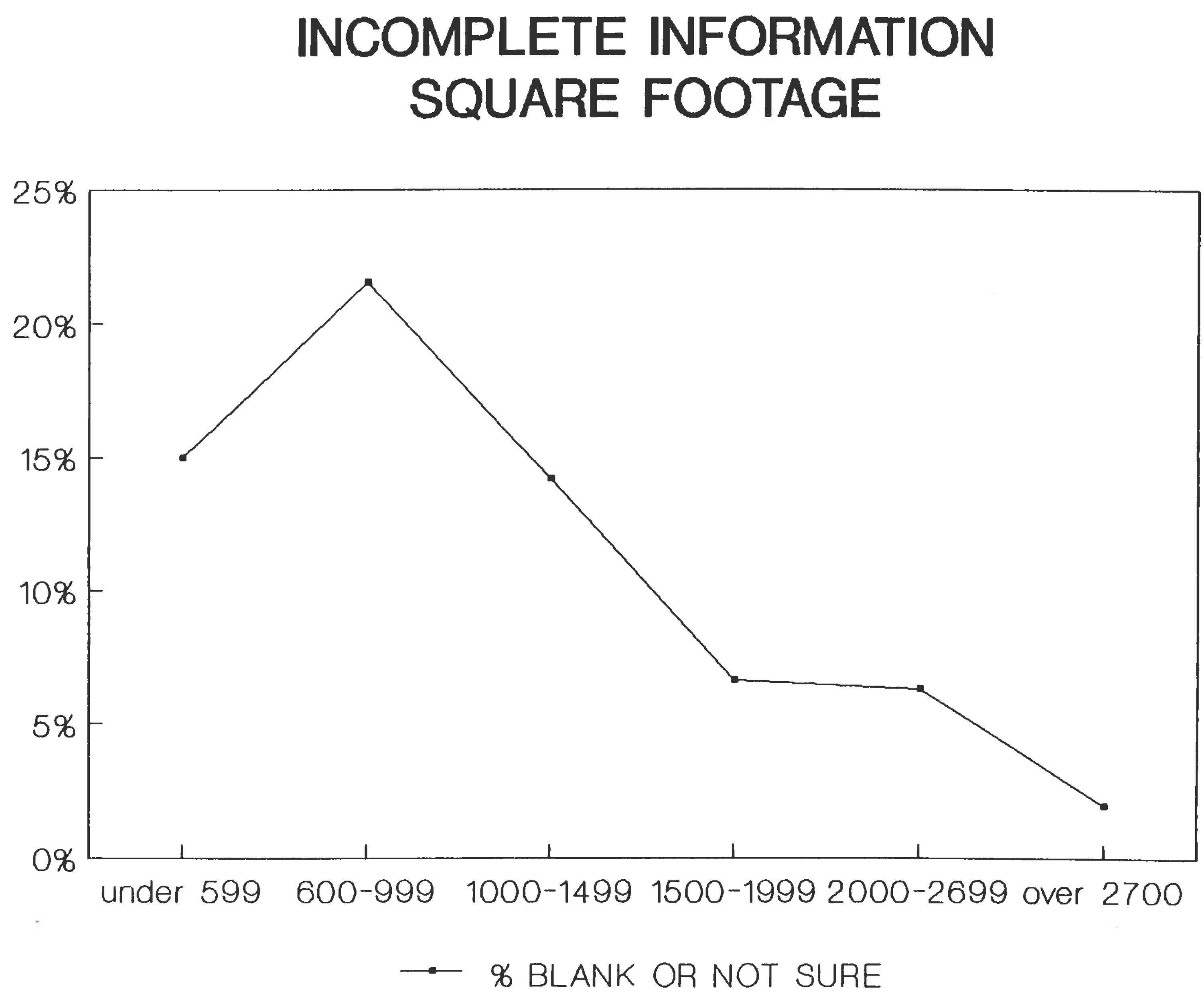

INCOMPLETE INFORMATION
SQUARE FOOTAGE
25%
20%
15%
10%
5%
0%
under 599
600-999
1000-1499
1500-1999
2000-2699
over 2700
% BLANK OR NOT SURE

Figure 3

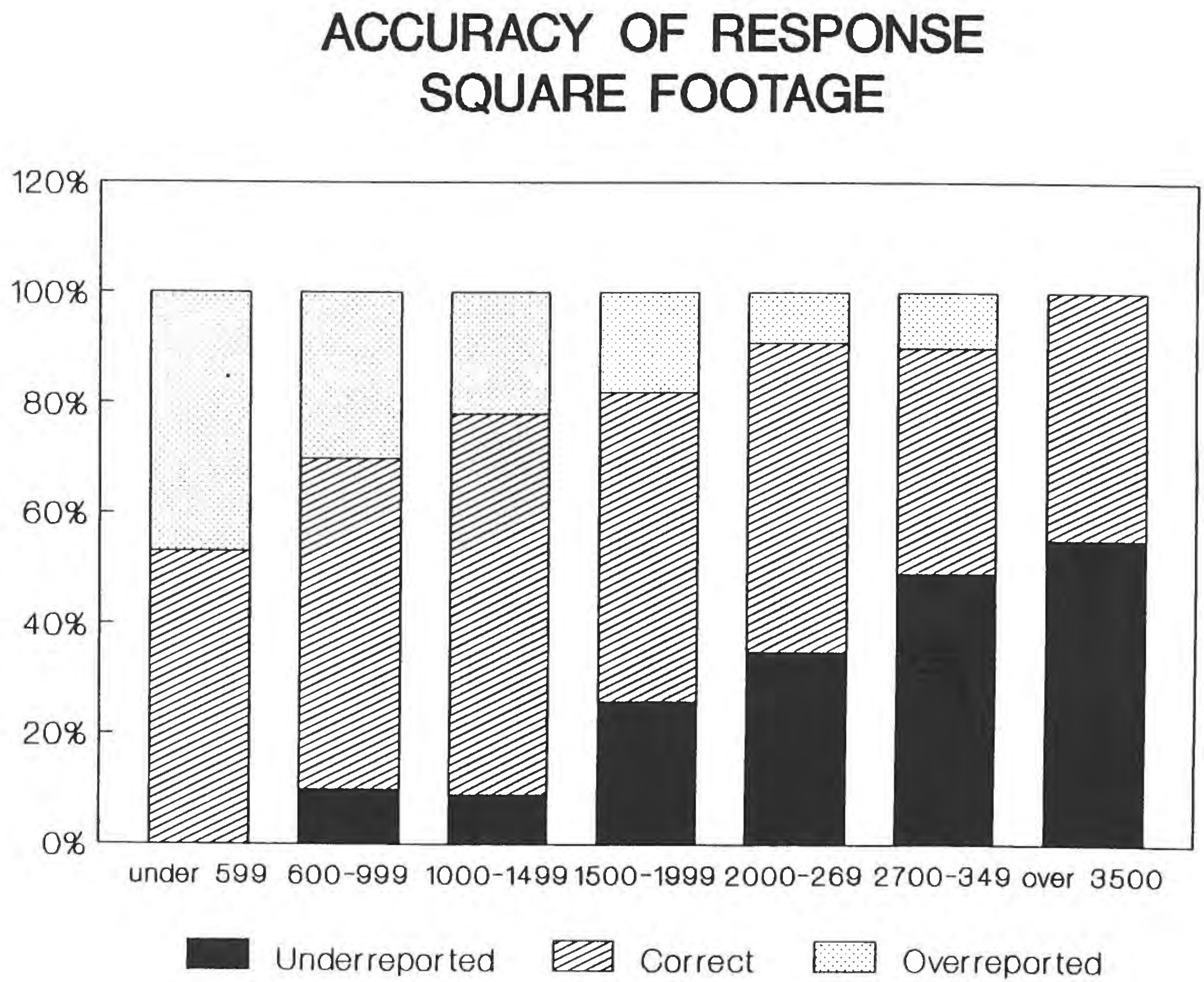

ACCURACY OF RESPONSE
SQUARE FOOTAGE
120%
100%
80%
60%
40%
20%
0%
under 599
600-999
1000-1499
1500-1999
2000-269
2700-349
over 3500
Underreported
Correct
Overreported

Figure 4

ESTIMATED SATURATION LEVELS BY TYPE OF THERMAL INSULATION

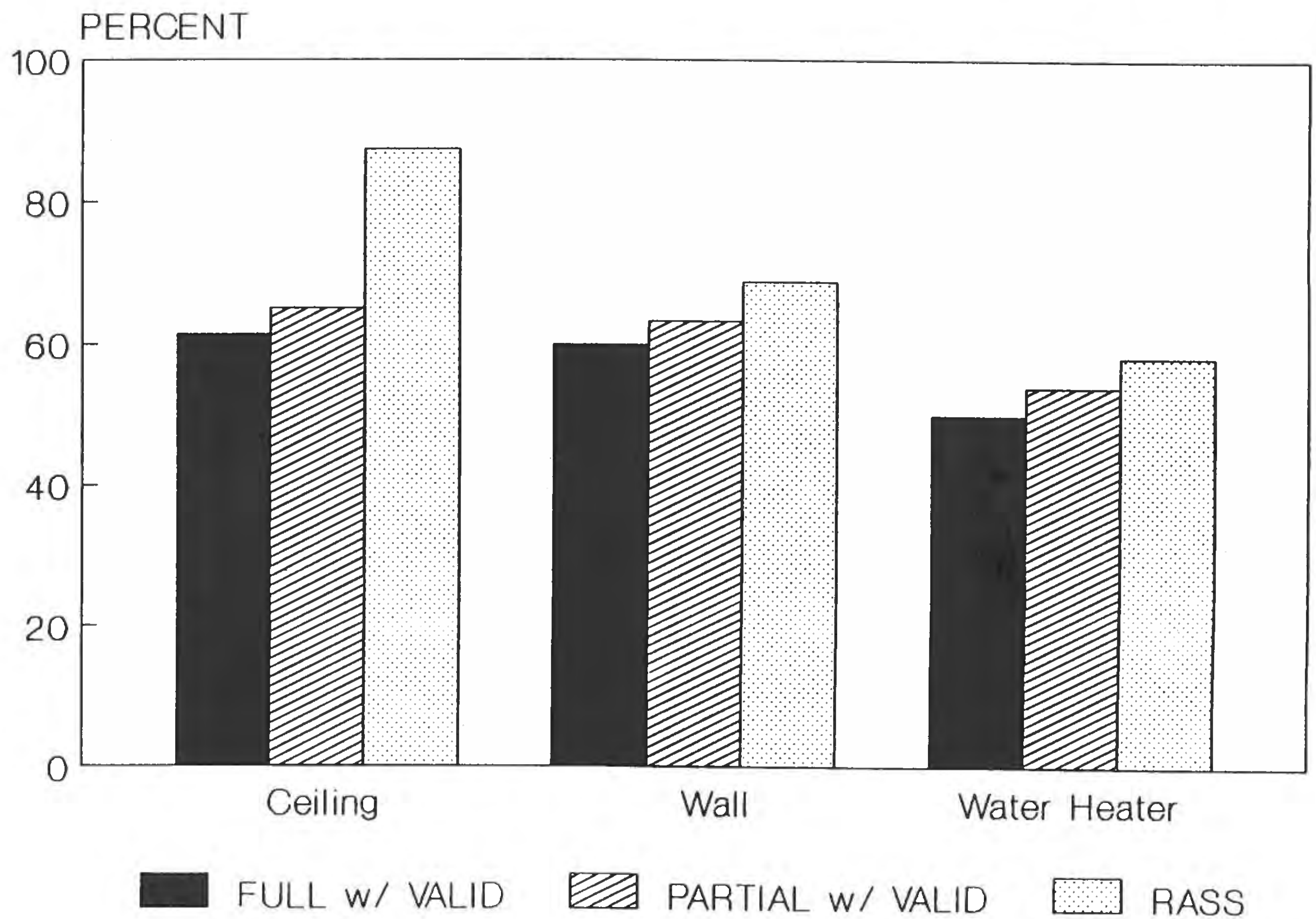


Figure 5

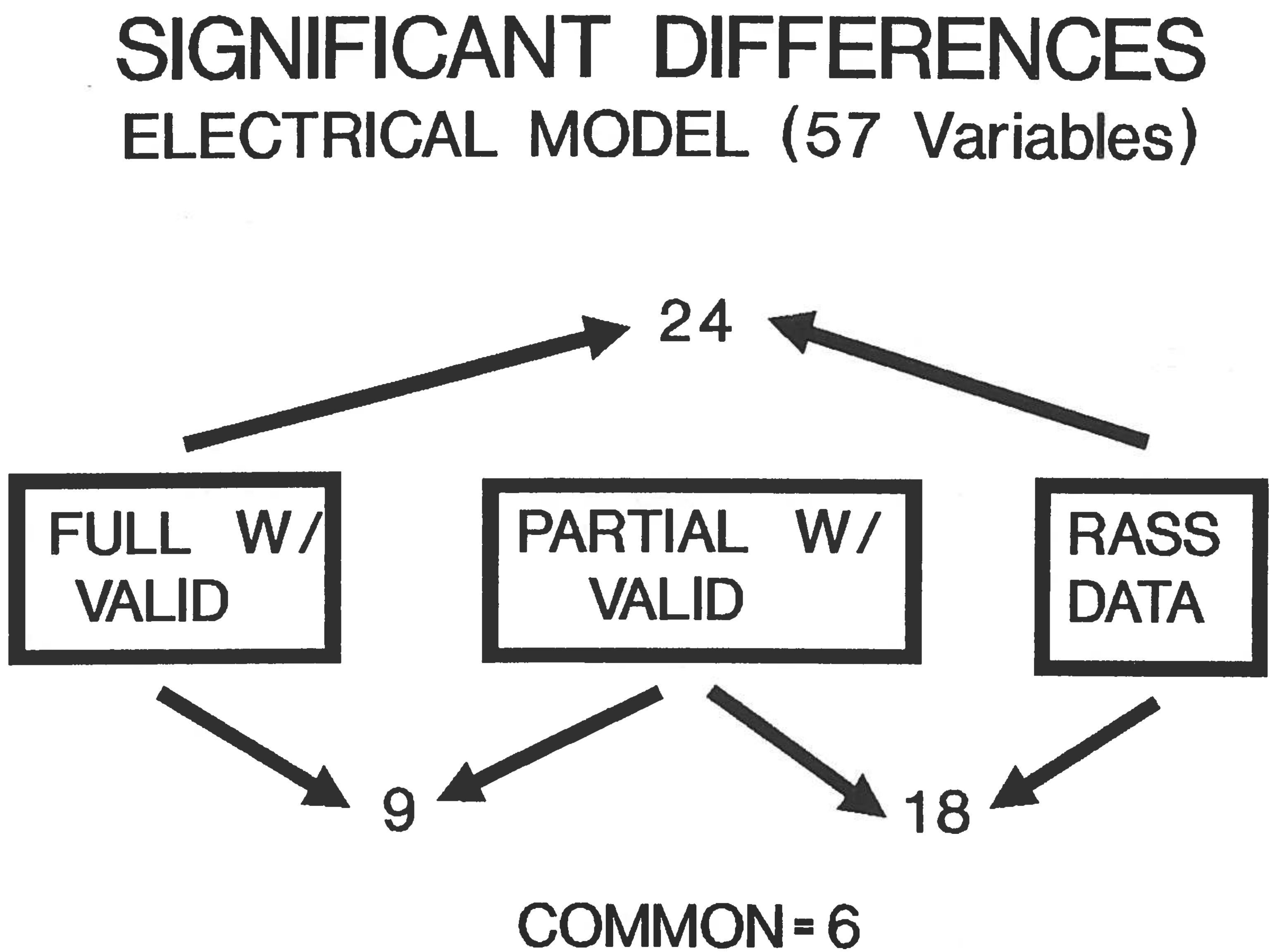
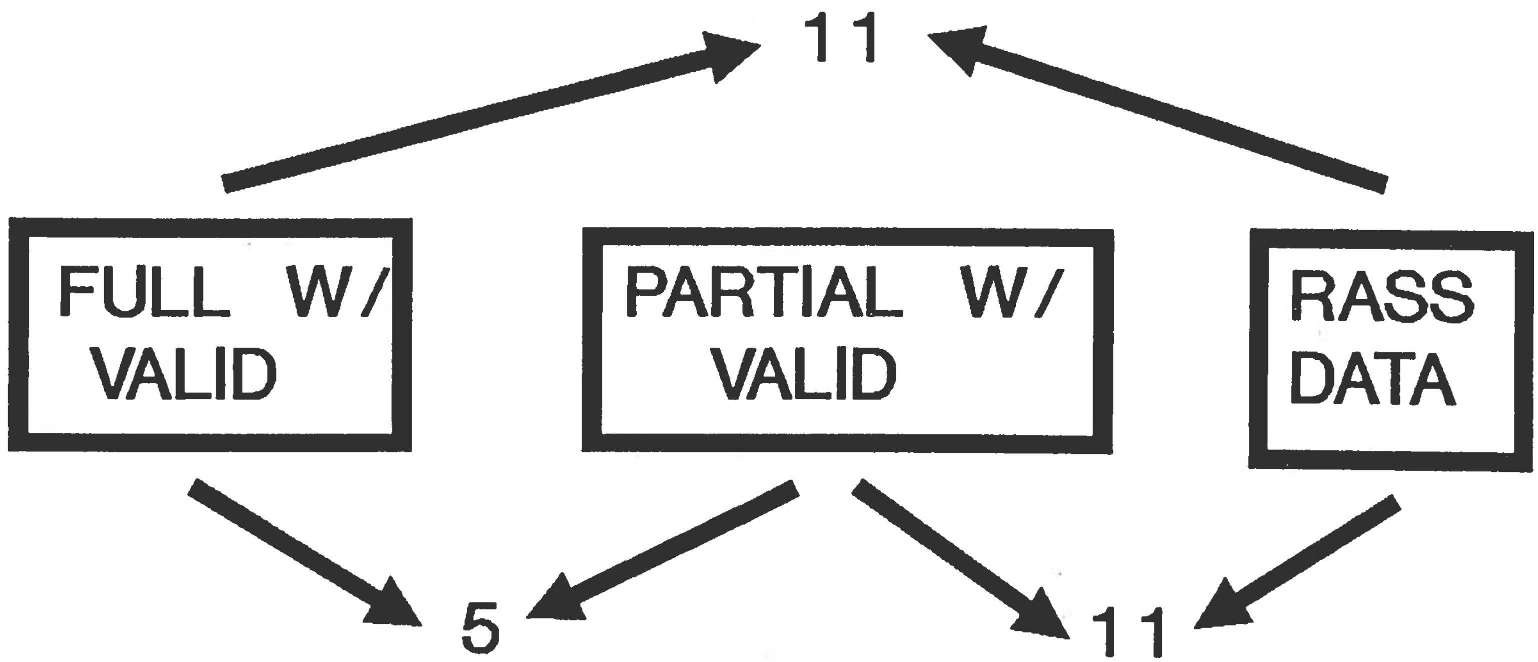

SIGNIFICANT DIFFERENCES
ELECTRICAL MODEL (57 Variables)
24
FULL W/ VALID
PARTIAL W/ VALID
RASS DATA
9
18
COMMON = 6

Figure 6

SIGNIFICANT DIFFERENCES
GAS MODEL (29 Variables)



COMMON = 3

Figure 7

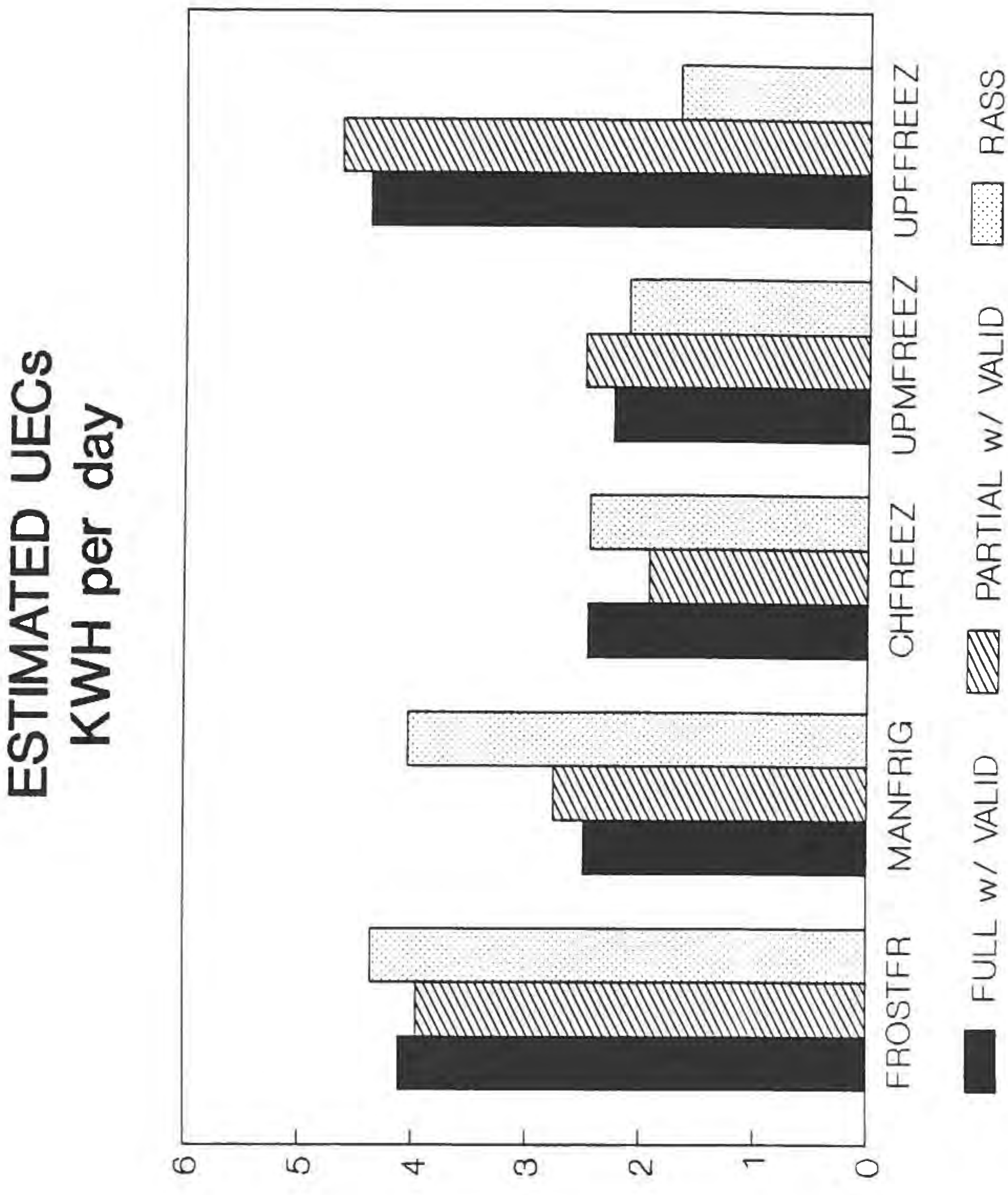

ESTIMATED UECs
KWH per day
0
1
2
3
4
5
6
FROSTFR
MANFRIG
CHFREEZ
UPMFREEZ
UPFFREEZ
FULL w/ VALID
PARTIAL w/ VALID
RASS

Figure 8

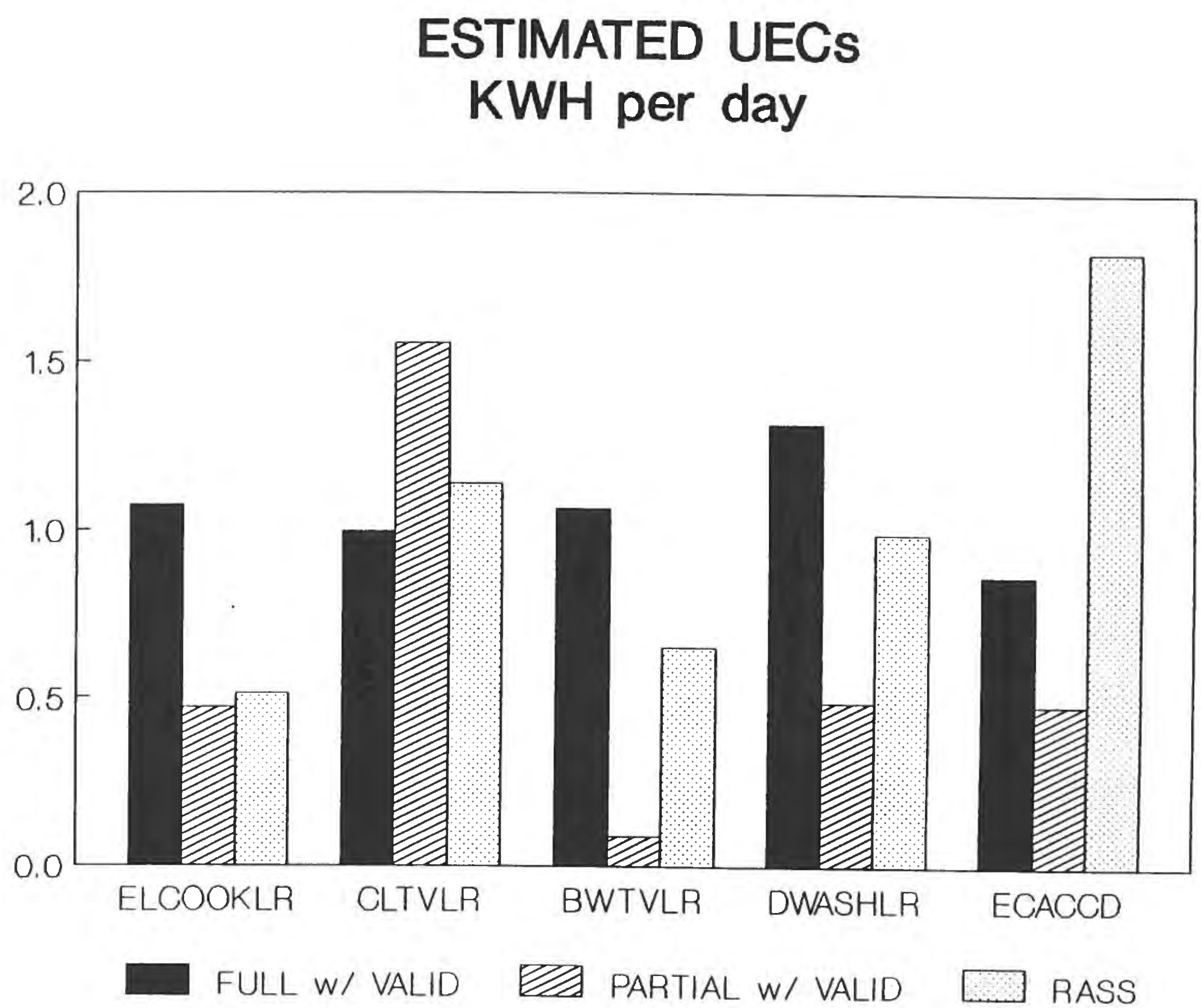


Figure 8 (contd.)