Title
AN ANALYSIS OF THE DIFFERENCES BETWEEN MONITORED INDOOR TEMPERATURES AND REPORTED THERMOSTAT SETTINGS

Authors
Vine, E.
Barnes, B.K.

Publication Date
1988-03-01

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

APPLIED SCIENCE DIVISION

An Analysis of the Differences between Monitored Indoor Temperatures and Reported Thermostat Settings

E. Vine and B.K. Barnes

March 1988

APPLIED SCIENCE DIVISION

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

DISCLAIMER

LBL-25035

AN ANALYSIS OF THE DIFFERENCES BETWEEN MONITORED INDOOR TEMPERATURES AND REPORTED THERMOSTAT SETTINGS

Edward Vine and B.K. Barnes

Energy Analysis Program
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

March 1988

Prepared for the
Office of Conservation and Power Resources
Bonneville Power Administration

This work was supported by the Bonneville Power Administration on Contract No. DE-AI79-84BP20521 through the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

ABSTRACT

We examined differences in reported winter thermostat settings and monitored temperatures, and contrasted those households with little difference, and those with a substantial difference. This analysis was conducted on households participating in Bonneville Power Administration's Residential Standards Demonstration Program (RSDP) in the Pacific Northwest. The reported thermostat settings were obtained from a survey of RSDP participants, and indoor temperatures were read from special recorders inside the house. We found reported thermostat settings to be on the average 2°F cooler than actual temperatures; differences between settings and temperatures were less for very energy-efficient homes than for homes built to current practice. Differences might be due to self-reporting inaccuracies. However, other explanations are also likely for the small difference (e.g., defective equipment, location of sensors, and physical processes not accounted for by thermostats (e.g., temperature gains from solar incidence, appliances, and occupants)), so that households might be accurately reporting their settings. This would then imply that, at the aggregate level, indoor temperature estimates can be calculated relatively inexpensively and accurately with the use of survey research techniques, in contrast to the more expensive approach of installing temperature sensors. However, we found substantial variation between self-reported thermostat settings and indoor temperatures at the household level. Accordingly, we recommend the use of temperature sensors for calculating heating and cooling loads for individual houses.

We contrasted those households with small differences (±2°F) between reported winter thermostat settings and monitored temperatures, and those with substantial differences (5°F or more). We were able to identify households with substantial differences based upon their space heating electricity use, the winter outdoor temperature, the physical/structural characteristics of the house, the number of appliances, the socioeconomic characteristics of the occupants, the energy behavior of the occupants, and/or energy-related attitudes. Using discriminant analysis, we were able to correctly classify 100% of the high-difference and low-difference groups using these variables.

A wide variety of sources of information are needed for discriminating between the high and low-difference groups. This information can be used by modelers and program managers for improving their energy-use and energy-saving estimates. For example, both the high and low-difference groups could be included in an analysis of energy use, but the first group's self-reported thermostat settings would be weighted (e.g., upwards by five degrees). Or the analysis of energy use would only be conducted for those households belonging to the low-difference group.

INTRODUCTION

Indoor temperatures are a key input in the modelling of household energy use. For example, space heating energy use is often viewed as highly dependent on the setting of the thermostat of the heater. Similarly, indoor temperatures are the focus of many residential energy conservation programs: low indoor temperatures during the winter and high indoor temperatures during the summer have been promoted for reducing energy use. However, the monitoring (metering) needed to obtain accurate data on indoor temperatures for modellers and program managers is very time-consuming and expensive, so that these individuals often rely on engineering assumptions, monitoring studies conducted in other parts of the country, or on self-reported thermostat setting data collected in household surveys. The usefulness of self-reported thermostat data as a surrogate for actual indoor temperatures is an open question and is one of the issues examined in this report.

Another issue explored in this paper is the analysis of households whose self-reported temperatures differ significantly from their metered temperatures (either higher or lower). Do these households have certain characteristics that distinguish them from households whose temperature differences are negligible? Can any of these characteristics be used effectively by program managers in designing and implementing programs by taking into account these differences in their calculations of energy use and energy savings? Can modellers use this information for adjusting their data inputs to improve their estimates and predictions?

We examine these issues by analyzing a group of homes in the Pacific Northwest which have been participating in the Bonneville Power Administration's Residential Standards Demonstration Program (RSDP) (Vine, 1986). The RSDP is a large-scale demonstration program of new, electrically-heated houses built to very energy-efficient standards (called the Model Conservation Standards, or MCS). Houses meeting the MCS are expected to use one-third to one-half of the heating energy of an otherwise comparable house built to current standards (the "current practice" house). The RSDP included the large-scale monitoring of both construction costs and energy use in approximately 400 energy-efficient houses and an equivalent number of "current practice" homes built recently to conventional standards (Vine, 1986; Meier et al, 1986). As part of the monitoring program, houses built to the MCS were "triple-metered" for electricity consumption by placing separate kilowatt-hour meters on the heating circuit, the domestic hot water circuit, and the total load. In addition, an integrating temperature recorder that measured both indoor and outdoor temperatures was installed. Cooperating homeowners were paid to record weekly the meter readings and indoor and outdoor temperatures.

Occupants of both the MCS and current practice houses were surveyed twice in the RSDP (March-May, 1985, and April-May, 1986). The second survey, the source of information used in this report, collected data on self-reported thermostat settings in the winter and in the summer for different periods of the day (see below). Information on house characteristics, appliances in the house, occupancy of the home, perceived problems with the indoor environment, energy-related attitudes, and demographic characteristics of the household was also obtained.

METHODOLOGY

In the occupant survey, respondents were asked to report their thermostat settings in winter (1985/86) under four different conditions. These conditions were (a) when people were at home and awake; (b) when people were asleep; (c) when nobody was at home during the day; and (d) when nobody was at home for more than a day. The monitored indoor temperatures were automatically averaged, so that our first task was to construct an "average" thermostat setting for the winter from the data we had collected. To do this, responses to two other questions in the occupant survey were considered. Respondents were asked what their house-occupancy profile was during the day (between 8 am and noon, noon and 4 pm, and 4 and 6 pm; for the remaining 14 hours, we assumed people were asleep for 8 hours and at home and awake for 6 hours). We also had some information concerning their weekend activities (if they were usually gone from their home during the weekend, their thermostat settings were assumed to be the same as when nobody was home during the day; if they were at home during the weekend, their thermostat settings were assumed to be the same as if they were at home during the weekday when they were asleep (8 hours) and at home and awake (16 hours)). From this information, and from the reported thermostat settings under the four conditions listed above, an estimate of the average reported thermostat setting was made. For the winter, if respondents reported that the thermostat was turned off at any time, a value was calculated for the thermostat setting, by subtracting 10°F from the lowest reported thermostat setting (see Vine and Barnes, 1986).

As part of the RSDP, temperatures in the home were monitored in both MCS and current practice houses, and average weekly indoor temperatures were read from the special recorder and reported by the household. For the winter season, we calculated the mean indoor temperature for the winter season (November 16, 1985 to February 15, 1986). It is important to note that the temperature sensors were located in a "blue box" containing the digital displays of the temperatures, so that some heat given off by this device was measured by the sensors. Accordingly, as part of the data collection process, adjustments of 2.4°F were made to all the temperature readings for all households to account for this external influence.† However, it is possible that the sensors were uniquely affected in each house, so that some error is introduced in the measured indoor temperatures analysed in this paper. In addition, about 12 RSDP homes had remote sensors located outside the blue box, so the temperature sensors were not affected by the energy used by the blue boxes in these houses. However, we did not separate these 12 houses from the rest of the sample in our analysis.

A COMPARISON OF INDOOR TEMPERATURES AND THERMOSTAT SETTINGS

Figure 1 presents the distributions of average monitored indoor winter temperatures and average reported winter thermostat settings. The mean monitored winter indoor temperature was 67.5°F with a standard deviation of 3.7°F. The mean reported winter thermostat setting was cooler: 65.2°F with a standard deviation of 4.5°F. The mean difference between monitored indoor temperatures and reported winter thermostat settings was 2.3°F with a standard deviation of 4°F.

Figure 2 presents the distribution of the differences between monitored temperatures and reported settings for MCS and current practice households. Previous studies of this RSDP sample (Vine and Barnes, 1986; Meier et al, 1986) have shown that MCS households have higher self-reported thermostat settings and monitored indoor temperatures than current practice households (statistically significant at the 0.05 level). Moreover, as shown in Figure 2, the difference between thermostat settings and indoor temperatures is lower for the MCS households (a mean of 0.9°F) than for current practice households (a mean of 2.3°F) (statistically significant at the 0.05 level). The differences for MCS and current practice households in each climate zone are presented in Figures 3 to 5; because of small sample sizes, it is difficult to say anything definitive about the effect of climate zone on differences between thermostat settings and monitored temperatures.

The mean difference (2.3°F) between indoor temperatures and thermostat settings may indicate that many households report their homes to be cooler in the winter than what they really are. However, other explanations for the differences may also be plausible. First, we expect indoor temperatures to rise (float) above thermostat settings because of the effects of solar gain and internal gains from appliances and occupants.

†Personal communication with Chuck Cramer, Lambert Engineering, July 29, 1987.

Figure 1. Comparison of Reported Thermostat Settings and Actual Indoor Temperature for Winter

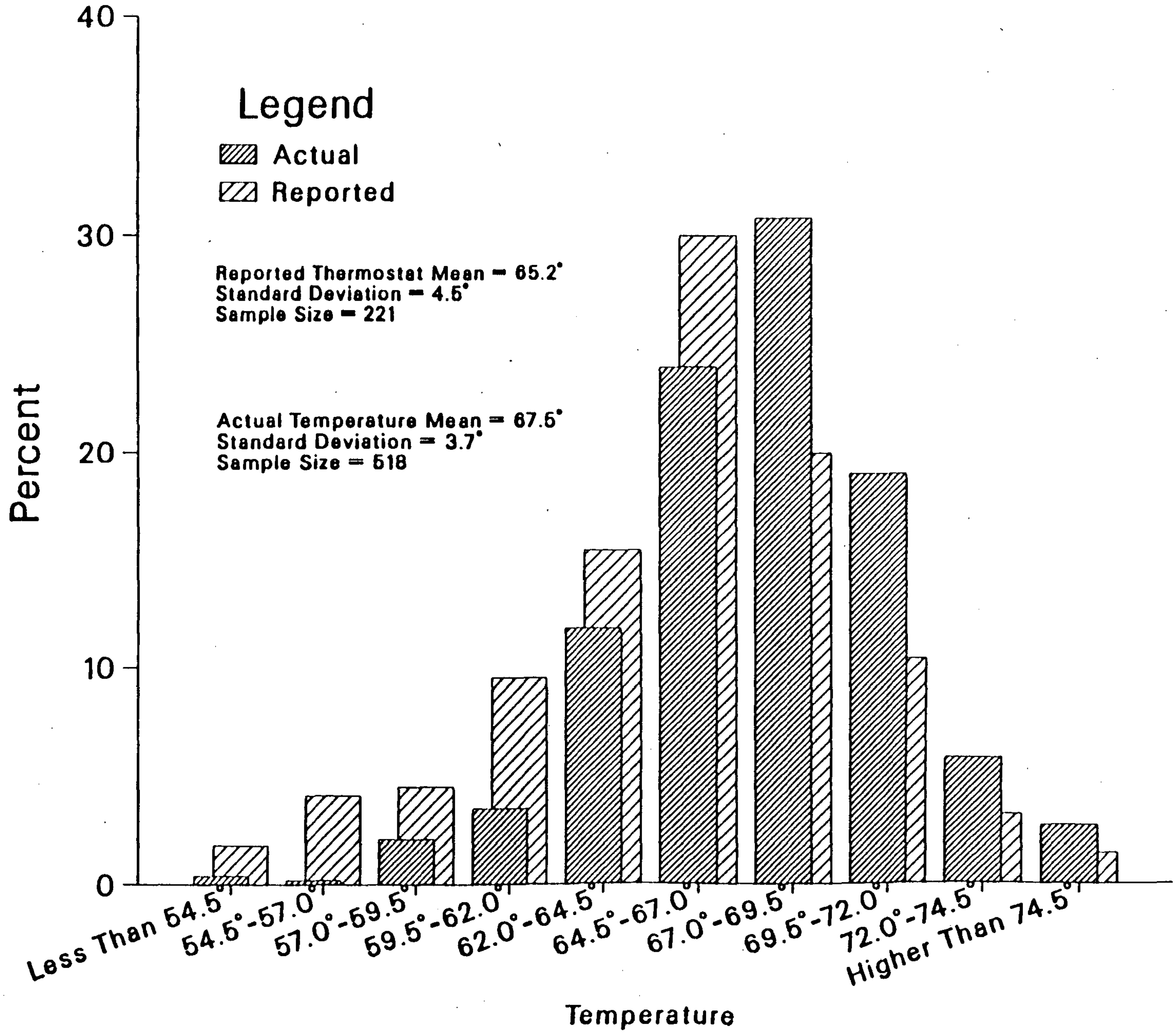


Figure 2. Difference Between Actual Indoor Temperatures and Reported Thermostat Settings for Winter

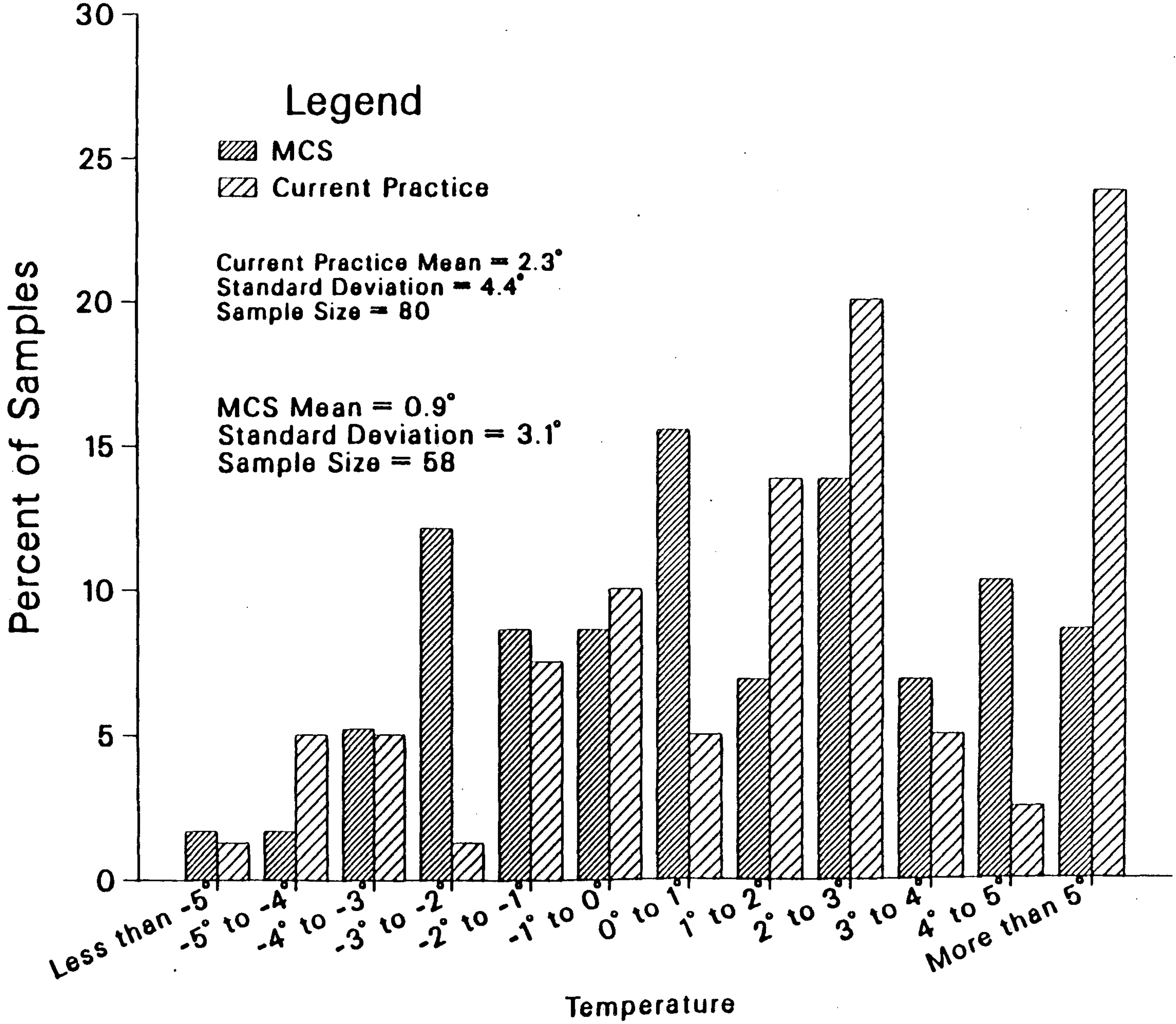


Figure 3. Difference Between Actual Indoor Temperatures and Reported Thermostat Settings for Winter Climate Zone 1

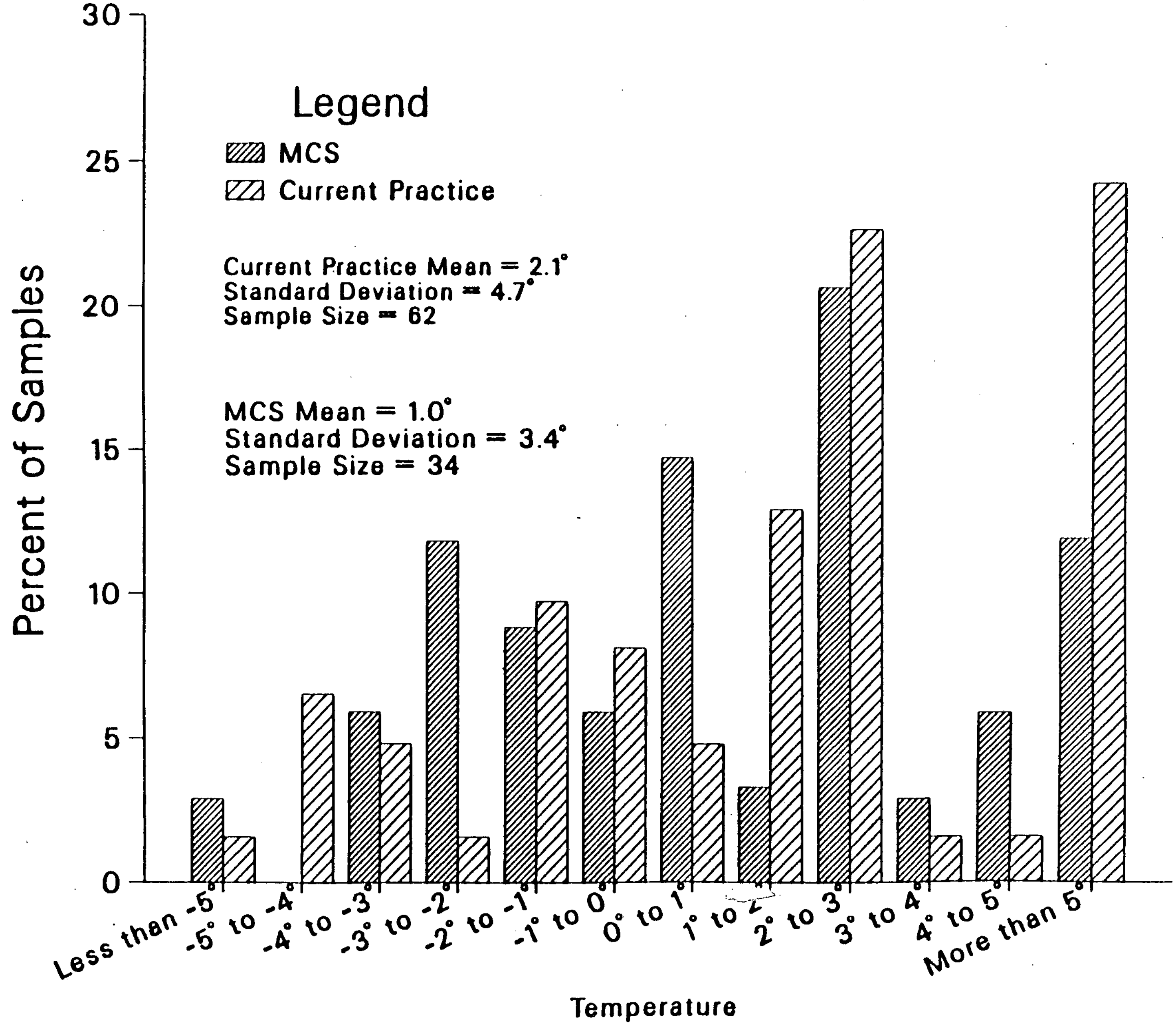
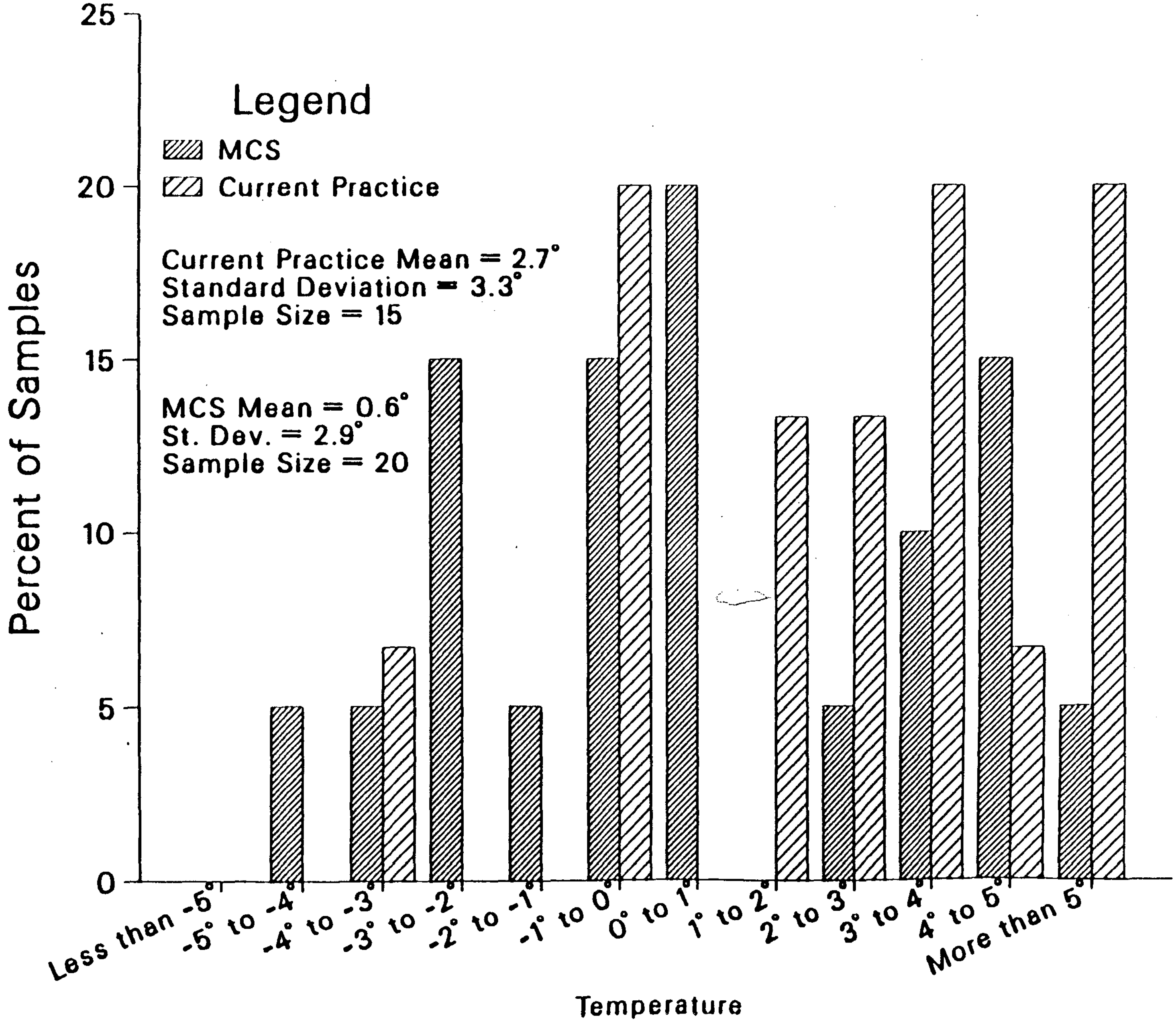
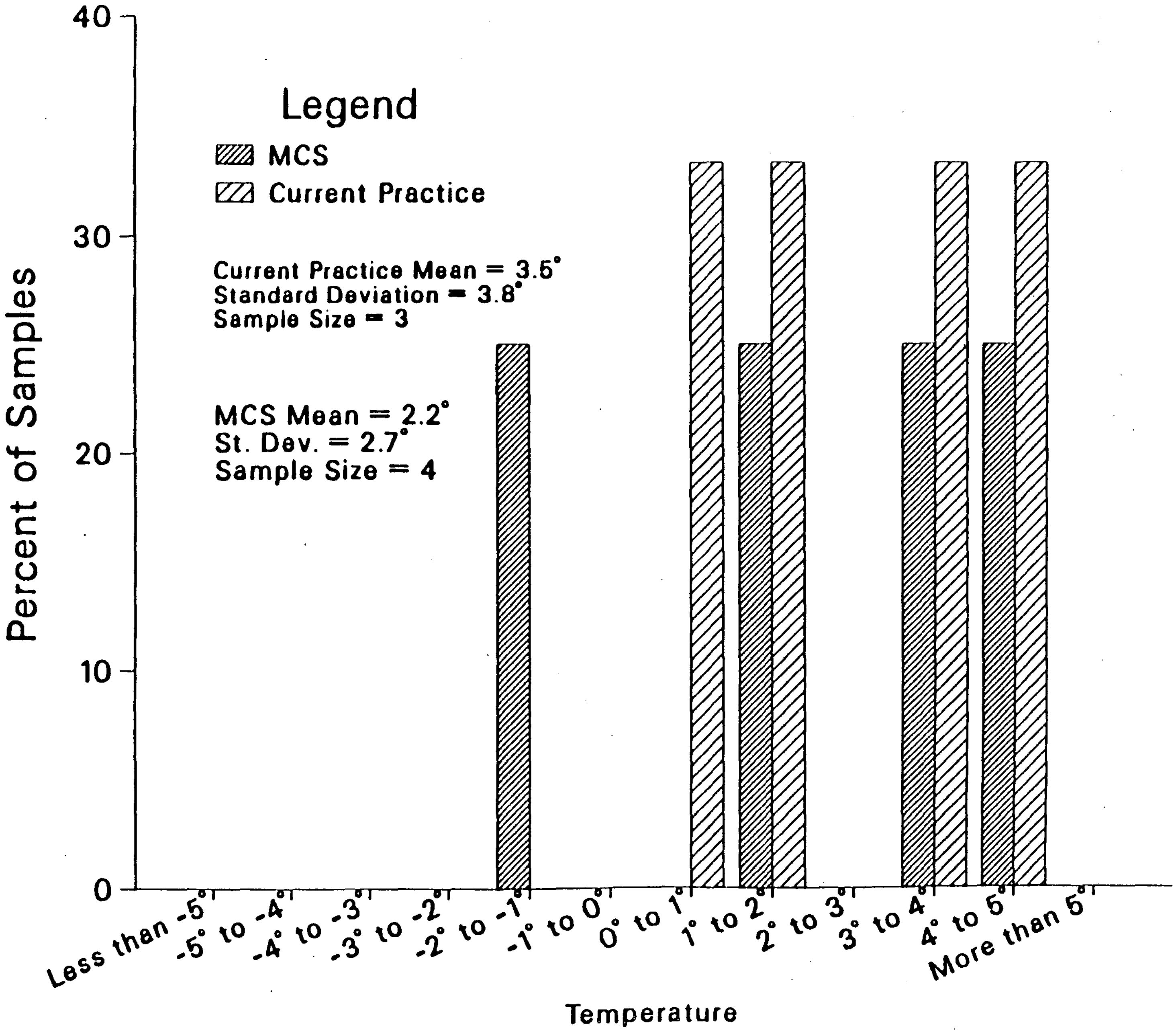


Figure 4. Difference Between Actual Indoor Temperatures and Reported Thermostat Settings for Winter Climate Zone 2



-6-

Figure 5. Difference Between Actual Indoor Temperatures and Reported Thermostat Settings for Winter Climate Zone 3



For example, on a sunny winter day, solar radiation will warm the interior of the building to temperatures above the thermostat set point. The 2.3 °F difference is relatively small and may reflect the influence of these gains, thereby indicating that the occupants are relatively accurate in reporting their indoor temperatures.

We should also mention some measurement-related problems that might cause differences between thermostat settings and indoor temperatures, so that self-reports might still be accurate. First, only a single internal temperature was taken in most of the RSDP homes. In most cases, this temperature was in the room directly controlled by the thermostat settings, for the single thermostat reported by the occupant; however, this may not be true in other cases. Second, even if the temperature sensors were all in the same rooms/zones as the reported thermostat setting, this temperature probably does not describe the average temperature in the house. Interior temperatures may vary from 5 to 10 °F between rooms (Palmiter and Hanford, 1986). Third, many of the RSDP homes had unitary heating systems (baseboard or wall-mounted forced air). These systems do not generally have calibrated thermostats, so reported thermostat settings may represent perceived or desired temperatures rather than control set points.[†] And fourth, the thermostats may not be operating properly: the settings on the device may not accurately reflect indoor temperatures, so we should expect some differences between thermostat settings and indoor temperatures.

DISCRIMINANT ANALYSIS

Because the difference between reported thermostat settings and indoor temperatures might be significant, we decided to compare households whose reported settings agreed with actual temperatures with those households whose settings and temperatures diverged. We found only two cases where reported thermostat settings were substantially larger (5 degrees or more) than measured temperatures. Therefore, we focused on the other outliers: where reported thermostat settings were less than measured temperatures (i.e., homes were warmer than what people reported them to be). We grouped those households with 5 °F or more into one group (Group H (the "high-difference group"): 23 households), and compared these households with those whose reported thermostat settings were within ±2 °F (Group L (the "low-difference group"): 53 households).[‡]

[†] Personal communication from Dick Byers, Research Analyst, Washington State Energy Office, August 18, 1987.

[‡] Members of Group H were randomly distributed across actual indoor temperatures and, therefore, were found in both warm and cool houses.

We used the statistical method of discriminant analysis to isolate those factors that distinguish Group H from households belonging to Group L. Discriminant analysis is an application of the general linear model and uses a linear combination of predictor variables to classify cases into one of two or more groups (Klecka, 1980). On the basis of such an analysis, it is possible to specify the variables which, in combination, can most effectly discriminate between the specified groups. It also provides a means of predicting group membership of cases where such membership is not already known. The details of our discriminant analysis are presented in the appendix, and we summarize the major findings of our analysis in this section.

Our approach in this analysis was exploratory and purely empirical: we wanted to look at all possible factors that might explain why a household would report lower thermostat settings than actually found in the house (or at least inferred from the indoor temperature data). We decided not to rely simply on physical, engineering models because previous research on energy use indicated that (1) household energy use can vary significantly in similarly, constructed buildings (Sonderegger, 1978), (2) the behavior of the occupants can play a very important role in affecting home energy use (Cramer et al, 1985; Seligman et al, 1981), and (3) existing models are typically able to explain only 40% to 50% of the variation in energy (e.g., Ritchie et al, 1981). Accordingly, given the wealth of available data at hand, we were eager to see which physical, behavioral, and attitudinal variables could best explain the temperature differences (monitored versus self-reported) in a given household.

The variables used in the discriminant analysis were based on data collected during the RSDP: builder cost forms, occupant surveys, and metering. Variables included space heating electricity use, outside winter temperature, location of house (state), type of heating system (central air, baseboard heat, heat pump, radiant heat), presence of a central thermostat, size of house (floor area), whether house was one-story or two-story, whether the house had a basement, the thermal integrity of the house (reported as an UA), structural modifications made to the house, number of appliances, problems with space conditioning equipment, heating-related questions (heating-related behavior (e.g., opening of windows and doors), thermal comfort, and performance of heating systems), perceptions of problems with indoor environment (mold/mildew, condensation, humidity, and odors), amount of time spent at home during the day and during the week, and demographics of occupants (household income, size, and age composition, and education, age, and sex of respondent).

Results

The first discriminant analysis was conducted on all households in order to distinguish households with large differences between self-reported thermostat settings and monitored indoor temperatures (Group H) from households with small differences (Group L). Thirty variables were included in the final model, broken down into six groups: energy use, weather, physical/structural, appliances, occupants, and behavior and attitudes (Table 1). None of the variables had strong correlations with the discriminating function (see Appendix). Eight variables were relatively important in discriminating between the two groups:

- number of electric space heaters
- winter outdoor temperature
- recently added wall insulation
- house size
- belief that it was easy to keep a comfortable temperature
- basement in home
- recently improved heating/cooling system
- central thermostat

Comparing the means of the two groups, households with more electric space heaters, recently added wall insulation, recently improved heating/cooling system, smaller houses, who disagreed that it was essential to have warm houses, and lived in warmer areas, were more likely to belong to Group H. On the other hand, households with basements in their houses and central thermostats were more likely to belong to Group L. This model was successful in explaining all (100%) of the grouped cases.

MCS households

We continued our use of discriminant analysis by treating MCS and current practice households separately. The discriminant function was first used to distinguish MCS Group H from MCS Group L. Nineteen variables were included in the final model, broken down into four groups: physical/structural, appliances, occupants, and behavior and attitudes (Table 2). We found seven variables to be particularly important in discriminating between the two groups:

- comparison of winter clothing to previous home
- recently improved heating/cooling system
- radiant heaters
- age of home
- central heater

Table 1. The Discriminant Model for All Households

Variables	Group H characteristics* (compared to Group L)
Energy use:	
Space heating electricity use	higher use
Weather:	
Winter outdoor temperature	warmer temperatures
Physical/Structural:	
House location (Oregon and Idaho)	less likely Oregon and Idaho
Size of home (floor area)	smaller homes
Age of home (year occupant(s) moved in)	older homes
Basement	less likely
Central thermostat present	less likely
Improved heating and cooling system	more likely
Added wall insulation	more likely
Added floor insulation	more likely
Appliances:	
Number of electric space heaters	more
Number of electric blankets	more
Number of dishwashers	less
Number of water heaters	less
Number of freezers	less
Number of second refrigerators	less
Number of televisions	more
Number of well pumps	more
Occupants:	
Size of household	smaller
Household income	wealthier
Babies in family	less likely
Age of respondent	older
Behavior and attitudes:	
Comparison of current winter thermostat setting with previous home setting	lower now
Home is heated quickly	more likely
Willing to wear heavy clothing	more strongly agree
House is stuffy or humid	more likely
Bedroom window kept open at night during winter	less likely
Hard to get rid of odors	more likely
Energy problem is not important	more strongly disagree

* For Group H households, reported thermostat settings were at least 5°F lower (cooler) than monitored indoor temperatures; for Group L households, reported thermostat settings were within ± 2°F of monitored indoor temperatures.

Table 2. The Discriminant Model for MCS Households

Variables	Group H characteristics (compared to Group L)*
Physical/Structural:	
Age of home (year occupant(s) moved in)	younger house
Two story house	more likely
Presence of fireplace or woodstove	less likely
Appliances:	
Presence of radiant heaters	more likely
Presence of baseboard heaters	less likely
Presence of central heater	less likely
Number of electric space heaters	more
Number of freezers	more
Number of well pumps	more
Occupants:	
Education of respondent	less educated
Age of respondent	younger
Household income	less wealthy
Behavior and attitudes:	
At-home during the day	less likely
Home unoccupied for more than 7 days	more likely
Improved heating/cooling system	more likely
Comparison of winter clothing to previous home	warmer now
Willing to wear heavy clothing	more strongly disagree
House is stuffy or humid	more likely
Energy problem is not important	more str. disagree

* For Group H households, reported thermostat settings were at least 5°F lower (cooler) than monitored indoor temperatures; for Group L households, reported thermostat settings were within ± 2°F of monitored indoor temperatures.

- number of well pumps[†]
- fireplace/woodstove

Comparing the means of the two groups, households who wore warmer clothes now than in their previous home, recently improved their heating/cooling system, had radiant heat, lived in more recently constructed homes, and had well pumps, were more likely to be members of Group H. In contrast, households with central heat and who had a fireplace or woodstove were more likely to be members of Group L. This model was successful in explaining 96% of the grouped cases: we were very successful for Group L households (100%), and somewhat less successful for Group H households (80%).

Current Practice households

The discriminant function was used to distinguish current practice Group H from current practice Group L. Twenty variables were included in the final model, broken down into five groups: energy use, physical/structural, appliances, occupants, and behavior and attitudes (Table 3). We found thirteen of these variables to be particularly important in discriminating between the two groups:

- recently installed weatherstripping
- recently added floor insulation
- recently added wall insulation
- recently added wood stove
- basement in house
- Montana house
- number of electric space heaters
- number of heat pump water heaters
- number of water heaters
- house is unoccupied for more than 7 days
- wanting to wear light clothing
- belief that it is essential to have a warm house
- age of respondent

Comparing the means of the two groups, households who recently installed weatherstripping, added floor and wall insulation, had electric space heaters, did not want to wear light clothing, did not believe it was essential to have a warm house, and had an older

[†] The presence of a well pump does not make theoretical sense since indoor temperatures are not affected by the presence of such equipment. More likely, well pumps are correlated with other variables not included in the model and, therefore, the relationship between well pumps and indoor temperatures is spurious.

Table 3. The Discriminant Model for Current Practice Households

Variables	Group H characteristics (compared to Group L)*
Energy use:	
Space heating electricity use	lower use
Physical/Structural:	
House location (Montana)	not in Montana
Age of home (year occupant(s) moved in)	same age
Basement	no basement
Installed weatherstripping	more likely
Added wall insulation	more likely
Added floor insulation	more likely
Added wood stove	less likely
Appliances:	
Number of electric space heaters	more
Number of electric blankets	more
Number of water heaters	less
Number of heat pump water heaters	none
Broken equipment	more likely
Occupants:	
Size of household	smaller
Age of respondent	older
Education of respondent	less educated
Behavior and attitudes:	
Essential to have house warm	more strongly disagree
Wanting to wear light clothing	more strongly disagree
Energy problem is not important	more strongly disagree
Home unoccupied for more than 7 days	less likely

* For Group H households, reported thermostat settings were at least 5°F lower (cooler) than monitored indoor temperatures; for Group L households, reported thermostat settings were within ± 2°F of monitored indoor temperatures.

person responding to the questionnaire, were more likely to be members of Group H. On the other hand, households who added a wood stove, had a basement in their house, lived in Montana, had water heaters, and did not occupy their house for 7 days or more at one time, were more likely to be members of Group L. This model was successful in explaining all (100%) of the grouped cases.

Reduced models

The model for all households was composed of 30 variables and was considered unwieldy. Accordingly, we decided to test a reduced form of the model by concentrating on a subset of the variables that would give us a satisfactory model. We ran ten models using subsets of variables taken from the complete analysis (Table 4 shows the results of the first four models). The first model included only one variable, the second model included two variables, the third three variables, and so on.

Table 4. Reduced Models for All Households

Variable	Model Number			
	1	2	3	4
Installed weatherstripping	x	x	x	x
Winter outdoor temperature		x	x	x
Number of electric space heaters			x	x
Presence of a central thermostat				x
Percent of Group H Correctly Classified	20.8	79.2	79.2	83.3

The measure of success for each model was its ability to successfully classify the high outliers. The model with four variables was able to classify 83% of the Group H households. Models with five or more variables did not improve our ability to correctly classify Group H households. Since the complete model was able to classify 100% of the cases, the model with four variables was relatively less accurate, however, still valuable in classifying Group H households when only using a few variables.

CONCLUSIONS AND DISCUSSION

In our analysis of households participating in Bonneville Power Administration's Residential Standards Demonstration Program, we examined differences in reported winter thermostat settings and monitored temperatures and found reported thermostat settings to be on the average 2°F cooler than actual temperatures, especially for current practice households, in contrast to MCS households. This difference might be due to self-reporting inaccuracies. However, other explanations are also likely (e.g., defective equipment, location of sensors, and physical processes not accounted for by thermostats (e.g., temperature gains from solar incidence, appliances, and occupants)), so that households might be accurately reporting their settings. This would then imply that, at the aggregate level, indoor temperature estimates can be calculated relatively inexpensively and accurately with the use of survey research techniques, in contrast to the more expensive approach of installing temperature sensors. However, we found substantial variation between self-reported thermostat settings and indoor temperatures at the household level. Accordingly, we recommend the use of temperature sensors for calculating heating and cooling loads for individual houses.

We also contrasted those households with small differences (+-2°F) between reported winter thermostat settings and monitored temperatures, and those with substantial differences (5°F or more). We were able to identify households with substantial differences based upon their space heating electricity use, the winter outdoor temperature, the physical/structural characteristics of the house, the number of appliances, the socioeconomic characteristics of the occupants, the energy behavior of the occupants, and/or energy-related attitudes. Using discriminant analysis, we were able to correctly classify 100% of the high-difference and low-difference groups using these variables.

These are, of course, optimistic estimates. Furthermore, because homeowners were paid to record the indoor and outdoor temperatures, they were likely to be more aware of their actual indoor temperatures at the time of the survey than other households. In other words, the "ecological validity" of the study is probably quite low, as it would be very difficult to generalize the findings from this sample to another sample.

We conducted discriminant analysis for three groups: all households, MCS households, and current practice households. In all three analyses, we found that households with electric space heaters and those who disagreed with the statement that the energy problem was not important were more likely to be high-difference households. Thus, in this case, as one explanation, differences between thermostat settings and indoor temperatures might be due to different heating practices using different heating systems, rather than inaccurate self-reporting (i.e., households might be accurately reporting their thermostat settings for their principal heating system, but also use auxillary heating systems that raise the indoor temperature).

In the analysis of all households, we found eight variables to be relatively important in discriminating between the two difference groups:

- number of electric space heaters
- winter outdoor temperature
- recently added wall insulation
- house size
- belief that it was easy to keep a comfortable temperature
- basement in home
- recently improved heating/cooling system
- central thermostat

Households with more electric space heaters, recently added wall insulation, recently improved heating/cooling system, smaller houses, who disagreed that it was essential to have warm houses, and lived in warmer areas, were more likely to belong to the high-difference group. On the other hand, households with basements in their houses and central thermostats were more likely to belong to the low-difference group.

The discriminant analysis indicated that the variables in the models were not highly correlated with the discriminant function, however, the models were very useful in correctly classifying the high and low-difference groups. Accordingly, a wide variety of sources of information are needed for discriminating between the two groups. Moreover, this information can be used by modelers and program managers for improving their energy-use and energy-saving estimates. For example, both the high and low-difference groups could be included in an analysis of energy use, but the first group's self-reported thermostat settings would be weighted (e.g., upwards by five degrees). Or the analysis of energy use would only be conducted for those households belonging to the low-difference group.

ACKNOWLEDGEMENTS

We would like to thank the following people for their assistance in this project: Ken Keating and Phil Thor of the Bonneville Power Administration; the participants in the RSDP; and Bruce Nordman and Alan Meier of Lawrence Berkeley Laboratory. We are also grateful for the comments provided by Dick Byers of the Washington State Energy Office on an earlier draft of this paper.

REFERENCES

Cramer, James C., Nancy Miller, Paul Craig, Bruce Hackett, Thomas Dietz, Edward Vine, Mark Levine, and Dan Kowalczyk, "Social and Engineering Determinants and Their Equity Implications in Residential Electricity Use." *Energy* 10(12):1283-91, 1985.

Klecka, William R., *Discriminant Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-019. Beverly Hills and London: Sage Publications, 1980.

Meier, Alan, Bruce Nordman, Craig Conner, and John Busch, *A Thermal Analysis of Homes in Bonneville Power Administration's Residential Standards Demonstration Program*. LBL Report 22109, Lawrence Berkeley Laboratory, Berkeley, Calif., 1986.

Palmiter, L. and J. Hanford, "Measured Performance of Superinsulated Homes in Montana: Second Report," *Proceedings of the 1986 Summer Study on Energy Efficiency in Buildings,* Amercian Council for an Energy Efficient Economy, Washington, D.C., 1986.

Ritchie, J.R., G. McDougall, and J. Claxton, "Complexities of Household Energy Consumption and Conservation." *Journal of Consumer Research* 8:233-242, 1981.

Seligman, C., L. Becker, and J.M. Darley, "Encouraging Residential Energy Conservation Through Feedback." In *Advances in Environmental Psychology, Vol. 3, Energy Conservation: Psychological Perspectives.* (Eds.) A. Baum and J.E. Singer. Lawrence Erlbaum Associates, Hillsdale, N.J., 1981.

Sonderegger, R. L., "Movers and Stayers: The Resident's Contribution to Variation Across Houses in Energy Consumption for Space Heating." In *Saving Energy in the Home: Princeton's Experiments at Twin Rivers.* (Ed.) R.H. Socolow. Ballinger, Cambridge, Mass., 1978.

Vine, Edward, *The Residential Standards Demonstration Program: Cost Analysis.* LBL Report 21318, Lawrence Berkeley Laboratory, Berkeley, Calif., 1986.

Vine, Edward and B.K. Barnes, *The Residential Standards Demonstration Program: Occupant Survey Analysis.* LBL Report 22409, Lawrence Berkeley Laboratory, Berkeley, Calif., 1986.

APPENDIX

In this appendix, we present in detail the results of the discriminant analysis used to isolate those factors that distinguish households with large differences between self-reported thermostat settings and monitored indoor temperatures (Group H) from households with small differences (Group L). We first present the results of the discriminant analysis for all the households in our sample, followed by the results for the MCS households, and then for the current practice households. Finally, we summarize the findings for a number of models based on subsets of variables used in the analysis of all the households.

Table A1a shows the mean values for the variables in the final discriminant model for the two groups when all households are analyzed. Comparing the means for the two groups is useful when interpreting the results of the discriminant analysis in the tables that follow.

We conducted the discriminant analysis in a stepwise fashion: variables were entered one at a time. In selecting variables for stepwise discriminant analysis, we used the minimization of **Wilks' lambda** as the criterion (at each step in introducing variables, the variable that results in the smallest Wilks' lambda for the discriminant function is selected for entry).† Table A1b shows the final stepwise discriminant model once the selection process was completed; 30 variables were included in the model. At each step in the model, the level of significance and Wilks' lambda indicate the utility of the function in classifying households into groups.‡ One expects the significance level of the model to increase (i.e., approach zero) and Wilks' lambda to decrease as variables are entered into the model, reflecting the improved ability of the model to classify households into groups.

For a useful discriminant function, the between-groups sum of squares should be greater than the within-groups sum of squares, and the ratio of the two is represented by

† The Wilks' lambda (sometimes called the U statistic) represents the ratio of within-groups sum of squares to the total sum of squares. In the two-group case, Wilks' lambda is the proportion of the total variance in the discriminant scores not explained by differences among groups. A lambda of 1 occurs when all observed group means are equal (i.e., there is no between-groups variability). Values close to 0 occur when within-groups variability is small compared to the total variability, that is, when most of the total variability is attributable to differences between the means of the groups. Thus, large values of lambda indicate that group means do not appear to be different, while small values indicate that group means do appear to be different.

‡ In Table A1b, variable Q2B1, which was entered in an earlier step, was removed from the model as other variables were entered and, essentially, negated the influence of the earlier variable. This action also occurred for some other variables in the model.

the **eigenvalue** of the function (large eigenvalues are associated with useful functions). The eigenvalue for the model was 14.92, indicating a useful function (Table A1c). In addition, the **canonical correlation** was very high (0.97).† The lambda also confirmed the relatively good discriminating value of this function, and the function was statistically significant.

To assess the contribution of a particular variable to the discriminating function, we relied on the correlation coefficients measuring the relationship between the discriminating variables and the discriminant function (Table A1d). The higher the correlation, the stronger the relationship between the variable and the function. Unfortunately, none of the variables had particularly strong correlations. Eight variables had correlation coefficients above 0.05:

- number of electric space heaters
- winter outdoor temperature
- recently added wall insulation
- house size
- belief that it was easy to keep a comfortable temperature
- basement in home
- recently improved heating/cooling system
- central thermostat

Comparing the means of the two groups (Table A1a), households with more electric space heaters, recently added wall insulation, recently improved heating/cooling system, smaller houses, who disagreed that it was essential to have warm houses, and lived in warmer areas, were more likely to belong to Group H. On the other hand, households with basements in their houses and central thermostats were more likely to belong to Group L.

The unstandardized discriminant function coefficients (Table A1e) are used to create the discriminant function to classify cases. This function was successful in predicting group membership in 100% of the grouped cases (Table A1f).

† The canonical correlation is a measure of the degree of association between the discriminant score and the group variable. In the two-group case, this value can be read as a simple Pearson correlation coefficient, ranging from 0 (no relationship) to 1 (identity).

Table A1. The Discriminant Model for All Households

A1a. Mean Values of Variables in Discriminant Model

	High Difference Group (H)	Low Difference Group (L)
(Q2B1) Installed Weatherstripping (1=no; 2=yes)	1.31	1.03
Winter Outdoor Temperature (°F)	35.18	30.38
(Q3J1) Number of Electric Space Heaters	1.81	0.76
Central Thermostat (0=no; 1=yes)	0.94	1.00
(Q2H1) Improved Heating and Cooling System (1=no; 2=yes)	1.06	1.00
(Q30A) House is Stuffy or Humid (1=no; 2=yes)	1.19	1.13
(Q12) Easy to Keep a Comfortable Temperature (1=yes; 3=no)	1.25	1.39
Size of Household	3.06	3.29
(Q2C1) Added Wall Insulation (1=no; 2=yes)	1.19	1.03
(Q17) Bedroom Window Kept Open in Winter (1=no; 3=yes)	1.00	1.13
Oregon (0=no; 1=yes)	0.25	0.26
(Q3G1) Number of Well Pumps	0.25	0.24
(Q42A) Broken Equipment (1=no; 2=yes)	1.25	1.26
(Q47A) Essential to Have House Warm (1=str. agree; 5=str. disagree)	2.12	1.63
(Q2A1) Added Rooms (1=no; 2=yes)	1.06	1.03
(Q31) Hard to Get Rid of Odors (1=yes; 2=no)	1.75	1.87
Basement (0=no; 1=yes)	0.06	0.29
(Space) Space Heating Electricity Use (kWh)	9612.00	9555.10
Idaho (0=no; 1=yes)	0.25	0.29
(Q3A1) Number of Dishwashers	0.88	0.95
(Q3E1) Number of Water Heaters	1.00	1.03
(Q47F) Energy Problem is Not Important (1=str. agree; 5=str. disagree)	4.19	4.09
(Q1A) Year Moved-in	82.94	83.45
Montana (0=no; 1=yes)	0.00	0.05
(Q16) Winter Thermostat Comparison (1=higher now; 3=lower now)	2.32	2.22
(Q11) Home is Heated Quickly (1=yes; 3=no)	1.25	1.45
Floor Area (square feet)	1755.94	2184.71
(Q2E1) Added Floor Insulation (1=no; 2=yes)	1.12	1.03
(Q3I1) Number of Electric Blankets	1.06	0.60
(Q3K1) Number of Televisions	1.75	1.71
(Q47E) Willing to Wear Heavy Clothing (1=str. agree; 5=str. disagree)	2.12	2.66
(Q48) Age of Respondent (1=younger than 18; 3=30-39; 6=65 or older)	3.69	3.50
(Q54) Household Income (1=less than $16,000; 6=$60,000 or more)	3.81	3.68
(Q3O1) Number of Freezers	0.62	0.74
(Q3N1) Number of Second Refrigerators	0.06	0.13
Babies in family (0=no; 1=yes)	0.50	0.55

A1b. Stepwise Discriminant Model for All Households

Step		Action Entered	Removed	Wilks' Lambda	Significance
1	(Q2B1) Installed Weatherstripping	1		0.83	0.00
2	Winter Outdoor Temperature	2		0.77	0.00
3	(Q3J1) Number of Electric Space Heaters	3		0.70	0.00
4	Central Thermostat	4		0.64	0.00
5	(Q2H1) Improved Heating and Cooling System	5		0.58	0.00
6	(Q30A) House is Stuffy or Humid	6		0.53	0.00
7	(Q12) Easy to Keep a Comfortable Temperature	7		0.48	0.00
8	Size of Household	8		0.45	0.00
9	(Q2C1) Added Wall Insulation	9		0.42	0.00
10	(Q17) Bedroom Window Kept Open	10		0.39	0.00
11	Oregon	11		0.37	0.00
12	(Q2B1) Installed Weatherstripping		10	0.37	0.00
13	(Q3G1) Number of Well Pumps	11		0.35	0.00
14	(Q42A) Broken Equipment	12		0.34	0.00
15	(Q47A) Essential to Have House Warm	13		0.32	0.00
16	(Q2A1) Added Rooms	14		0.31	0.00
17	(Q31) Hard to Get Rid of Odors	15		0.29	0.00
18	Basement	16		0.28	0.00
19	(Space) Space Heating Electricity Use	17		0.26	0.00
20	Idaho	18		0.25	0.00
21	(Q3A1) Number of Dishwashers	19		0.24	0.00
22	(Q3E1) Number of Water Heaters	20		0.22	0.00
23	(Q47F) Energy Problem is Not Important	21		0.21	0.00
24	Oregon		20	0.21	0.00
25	(Q1A) Year Moved-in	21		0.20	0.00
26	(Q47A) Essential to Have House Warm		20	0.21	0.00
27	(Q42A) Broken Equipment		19	0.21	0.00
28	Montana	20		0.20	0.00
29	(Q16) Winter Thermostat Comparison	21		0.19	0.00
30	(Q11) Home is Heated Quickly	22		0.16	0.00
31	Floor Area	23		0.14	0.00
32	(Q12) Easy to Keep a Comfortable Temperature		22	0.15	0.00
33	(Q2E1) Added Floor Insulation	23		0.13	0.00
34	(Q2A1) Added Rooms		22	0.13	0.00
35	(Q3I1) Number of Electric Blankets	23		0.12	0.00
36	(Q3K1) Number of Televisions	24		0.11	0.00
37	(Q47E) Willing to Wear Heavy Clothing	25		0.09	0.00
38	Oregon	26		0.09	0.00
39	(Q48) Age of Respondent	27		0.08	0.00
40	(Q54) Household Income	28		0.07	0.00
41	(Q3O1) Number of Freezers	29		0.07	0.00
42	Montana		28	0.07	0.00
43	(Q3N1) Number of Second Refrigerators	29		0.07	0.00
44	Babies in family	30		0.06	0.00

A1c. Canonical Discriminant Function

Eigenvalue	Canonical Correlation	Wilks' Lambda	Chi-Squared	Degrees of Freedom	Significance
14.92	0.97	0.06	102.40	30	0.00

A1d. Correlations Between Discriminating Variables and Canonical Discriminant Function

Q3J1	0.09
Winter Outdoor Temperature	0.08
Q2C1	0.08
Floor Area	-0.07
Q12	-0.07
Basement	-0.07
Q2H1	0.06
Central Thermostat	-0.06
Q47E	-0.05
Q2E1	0.05
Q17	-0.04
Q3I1	0.04
Q11	-0.04
Q31	-0.04
Q3A1	-0.03
Q1A	-0.03
Q3N1	-0.03
Q48	0.02
Size of Household	-0.02
Q16	0.02
Q30A	0.02
Q47F	0.01
Q54	0.01
Q3E1	-0.01
Babies in family	-0.01
Idaho	-0.01
Q3K1	0.01
Oregon	-0.004
Q3G1	0.003
Space	0.001

A1e. Canonical Discriminant Function Coefficients

	Unstandardized	Standardized
Q1A	0.73	1.45
Q2C1	2.26	0.58
Q2E1	6.83	1.56
Q2H1	7.90	1.06
Q3A1	16.98	4.49
Q3E1	-4.82	-1.15
Q3G1	3.98	1.90
Q3I1	0.62	0.81
Q3J1	2.58	3.64
Q3K1	1.64	1.30
Q3N1	1.38	0.44
Q3O1	-1.12	-0.67
Q11	-4.70	-2.74
Q16	-3.30	-1.74
Q17	-2.22	-0.77
Q30A	2.38	0.86
Q31	-7.77	-2.92
Q47E	0.47	0.56
Q47F	-1.24	-1.13
Q48	1.07	1.08
Q54	0.77	0.92
Size of Household	1.33	1.64
Babies in Family	0.93	0.47
Idaho	5.45	2.48
Oregon	-1.57	-0.70
Winter Outdoor Temperature	0.41	2.99
Basement	-0.92	-0.38
Central Thermostat	-29.34	-3.94
Floor Area	0.001	0.58
Space	0.001	3.09

A1f. Classification Results

Actual Group	Number of Cases	Predicted Group Membership H	L
Group H	16	16 100.0%	0 0.0%
Group L	38	0 0.0%	38 100.0%

Percent of Grouped Cases Correctly Classified: 100.0%

MCS high outliers

We continued our use of discriminant analysis by treating MCS and current practice households separately. The discriminant function was first used to distinguish MCS Group H from MCS Group L.

Table A2a shows the mean values for the variables in the final discriminant model for the MCS households. Table A2b shows the final stepwise discriminant model; 19 variables were included in the model. The eigenvalue of the function was 2858, indicating a very useful function (Table A2c). The canonical correlation was also very high (0.99), and the function was statistically significant.

The correlation coefficients for the discriminating variables and the discriminant function again indicated weak correlations (Table A2d). Seven variables had correlations of 0.01 or above:

- comparison of winter clothing to previous home
- recently improved heating/cooling system
- radiant heaters
- age of home
- number of well pumps
- central heater
- fireplace/woodstove

Comparing the means of the two groups (Table A2a), households who wore warmer clothes now than in their previous home, recently improved their heating/cooling system, had radiant heat, lived in more recently constructed homes, and had well pumps, were more likely to be members of Group H. In contrast, households with central heat and who had a fireplace or woodstove were more likely to be members of Group L. The discriminant function was successful in predicting group membership in 96% of the grouped MCS cases (Table A2f).

Table A2. The Discriminant Model for MCS Households

A2a. Mean Values of Variables in Discriminant Model

	High Difference Group (H)	Low Difference Group (L)
(Q15B) Winter Clothing Comparison (1=warmer now; 3=lighter now)	1.50	1.94
Radiant Heaters (0=no; 1=yes)	0.25	0.00
(Q3G1) Number of Well Pumps	0.75	0.23
(Q2H1) Improved Heating/Cooling System (1=no; 2=yes)	1.25	1.00
(Q25A) Home Unoccupied More Than 7 Days (1=no; 2=don't know; 3=yes)	2.25	1.64
(Q48) Age of Respondent (1=younger than 18;3=30-39;6=65 or older)	3.50	3.54
(Q47F) Energy Problem is Not Important (1=str. agree; 5=str. disagree)	4.50	4.14
Number of Stories (0=two story; 1=one story)	0.25	0.45
Fireplace/woodstove (0=no; 1=yes)	0.25	0.64
Floor Area (square feet)	2005.50	2477.09
(Q1A) Year Moved-in	85.00	84.33
At-home During the Day (0=no; 1=yes)	0.75	0.77
(Q3J1) Number of Electric Space Heaters	0.75	0.64
Baseboard Heaters (0=no; 1=yes)	0.00	0.04
Central Heater (0=no; 1=yes)	0.00	0.45
(Q47E) Willing to Wear Heavy Clothing (1=str. agree; 5=str. disagree)	3.00	2.95
(Q30A) House is Stuffy or Humid (1=no; 2=yes)	1.25	1.18
(Q54) Household Income (1=less than $16,000; 6=$60,000 or more)	3.75	3.90
(Q3O1) Number of Freezers	0.75	0.59
(Q53) Education of Respondent (low to high)	3.50	3.54

A2b. Stepwise Discriminant Model for MCS Households

Step		Action Entered	Removed	Wilks' Lambda	Significance
1	(Q15B) Winter Clothing Comparison	1		0.75	0.01
2	Radiant Heaters	2		0.49	0.00
3	(Q3G1) Number of Well Pumps	3		0.32	0.00
4	(Q2H1) Improved Heating/Cooling System	4		0.23	0.00
5	(Q25A) Home Unoccupied More Than 7 Days	5		0.19	0.00
6	(Q48) Age of Respondent	6		0.15	0.00
7	(Q47F) Energy Problem Is Not Important	7		0.11	0.00
8	Number of Stories	8		0.09	0.00
9	Fireplace/woodstove	9		0.07	0.00
10	Floor Area	10		0.06	0.00
11	(Q1A) Year Moved-in	11		0.05	0.00
12	At-home During the Day	12		0.03	0.00
13	(Q3J1) Number of Electric Space Heaters	13		0.02	0.00
14	Baseboard Heaters	14		0.01	0.00
15	Floor Area		13	0.01	0.00
16	Central Heater	14		0.01	0.00
17	(Q47E) Willing to Wear Heavy Clothing	15		0.003	0.00
18	(Q30A) House is Stuffy or Humid	16		0.002	0.00
19	(Q54) Household Income	17		0.001	0.00
20	(Q3O1) Number of Freezers	18		0.0005	0.00
21	(Q53) Education of Respondent	19		0.0004	0.00

A2c. Canonical Discriminant Function

Eigenvalue	Canonical Correlation	Wilks' Lambda	Chi-Squared	Degrees of Freedom	Significance
2858.26	0.99	0.0003	115.40	19	0.00

A2d. Correlations Between Discriminating Variables and Canonical Discriminant Function

Q15B	-0.01
Q2H1	0.01
Radiant Heaters	0.01
Q1A	0.01
Q3G1	0.01
Central Heater	-0.01
Fireplace/woodstove	-0.01
Q25A	0.005
Number of Stories	-0.003
Q47F	0.002
Q3O1	0.002
Q3J1	0.002
Baseboard Heaters	-0.002
Q30A	0.001
Q54	-0.001
At-home	-0.0004
Q53	-0.0003
Q48	-0.0003
Q47E	0.0003

A2e. Canonical Discriminant Function Coefficients

	Unstandardized	Standardized
Q1A	26.74	15.55
Q2H1	45.03	7.96
Q3G1	24.85	13.04
Q3J1	-30.98	-15.28
Q3O1	-4.33	-2.51
Q15B	-105.44	-30.29
Q25A	17.39	15.82
Q30A	-13.20	-5.40
Q47E	4.43	5.34
Q47F	-14.14	-14.61
Q48	12.26	13.81
Q53	-0.78	-0.85
Q54	-4.66	-4.46
Fireplace/woodstove	-34.14	-16.84
At-home	21.47	9.41
Baseboard Heaters	20.72	4.13
Central Heater	-4.11	-1.96
Radiant Heaters	149.74	26.47
Number of Stories	-10.53	-5.35

A2f. Classification Results

Actual Group	Number of Cases	Predicted Group Membership H	L
Group H	5	4 80.0%	1 20.0%
Group L	23	0 0.0%	23 100.0%

Percent of Grouped Cases Correctly Classified: 96.4%

Current practice high outliers

The discriminant function was used to distinguish current practice Group H households from current practice Group L households. Table A3a shows the mean values for the variables in the final discriminant model for the current practice households. Table A3b shows the final stepwise discriminant model; 20 variables were included in the model. The eigenvalue of the function was 900, indicating a very useful function (Table A3c). The canonical correlation was also very high (0.99), and the function was statistically significant.

The correlation coefficients for the discriminating variables and the discriminant function were very low (Table A3d). Thirteen of these variables had correlation coefficients of 0.01 or above:

- recently installed weatherstripping
- recently added floor insulation
- recently added wall insulation
- recently added wood stove
- basement in house
- Montana house
- number of electric space heaters
- number of heat pump water heaters
- number of water heaters
- house is unoccupied for more than 7 days
- wanting to wear light clothing
- belief that it is essential to have a warm house
- age of respondent

Comparing the means of the two groups (Table A3a), households who recently installed weatherstripping, added floor and wall insulation, had electric space heaters, did not want to wear light clothing, did not believe it was essential to have a warm house, and had an older person responding to the questionnaire, were more likely to be members of Group H. On the other hand, households who added a wood stove, had a basement in their house, lived in Montana, had water heaters, and did not occupy their house for 7 days or more at one time, were more likely to be members of Group L. The discriminant function was successful in predicting group membership in 100% of the grouped current practice cases (Table A3f).

Table A3. The Discriminant Model for Current Practice Households

A3a. Mean Values of Variables in Discriminant Model

	High Difference Group (H)	Low Difference Group (L)
(Q2B1) Installed Weatherstripping (1=no; 2=yes)	1.42	1.06
Basement (0=no; 1=yes)	0.00	0.31
(Q3I1) Number of Electric Blankets	1.25	0.81
(Q17) Bedroom Window Kept Open in Winter (1=no; 3=yes)	1.00	1.25
(Q1A) Year Moved-in	82.26	82.25
(Q2C1) Added Wall Insulation (1=no; 2=yes)	1.25	1.06
(Q3F1) Number of Heat Pump Water Heaters	0.00	0.06
(Q3J1) Number of Electric Space Heaters	2.17	0.94
(Q48) Age of Respondent (1=younger than 18; 3=30-39; 6=65 or older)	3.75	3.44
(Q2G1) Added Wood Stove	1.00	1.12
Size of Household	2.92	3.19
(Space) Space Heating Electricity Use (kWh)	9884.00	10910.25
(Q47E) Willing to Wear Heavy Clothing (1=str. agree; 5=str. disagree)	1.83	2.25
(Q53) Education of Respondent (low to high)	3.88	4.04
(Q47F) Energy Problem is Not Important (1=str. agree; 5=str. disagree)	4.08	4.00
Montana (0=no; 1=yes)	0.00	0.06
(Q47A) Essential to Have House Warm (1=str. agree; 5=str. disagree)	2.25	1.81
(Q25A) Home Unoccupied More Than 7 Days (1=no; 2=don't know; 3=yes)	1.50	1.94
(Q3E1) Number of Water Heaters	1.00	1.06
(Q42A) Broken Equipment (1=no; 2=yes)	1.25	1.19
(Q2E1) Added Floor Insulation (1=no; 2=yes)	1.17	1.06
(Q47B) Want to Wear Light Clothing (1=str. agree; 5=str. disagree)	4.17	3.50

A3b. Stepwise Discriminant Model for Current Practice Households

Step		Action Entered	Removed	Wilks' Lambda	Significance
1	(Q2B1) Installed Weatherstripping	1		0.82	0.02
2	Basement	2		0.72	0.02
3	(Q3I1) Number of Electric Blankets	3		0.64	0.01
4	(Q17) Bedroom Window Kept Open in Winter	4		0.57	0.01
5	(Q1A) Year Moved-in	5		0.48	0.00
6	(Q2C1) Added Wall Insulation	6		0.38	0.00
7	(Q3F1) Number of Heat Pump Water Heaters	7		0.31	0.00
8	(Q3J1) Number of Electric Space Heaters	8		0.27	0.00
9	(Q48) Age of Respondent	9		0.22	0.00
10	(Q2G1) Added Wood Stove	10		0.18	0.00
11	Size of Household	11		0.15	0.00
12	(Space) Space Heating Electricity Use	12		0.10	0.00
13	(Q47E) Willing to Wear Heavy Clothing	13		0.07	0.00
14	(Q53) Education of Respondent	14		0.05	0.00
15	(Q47F) Energy Problem is Not Important	15		0.04	0.00
16	Montana	16		0.02	0.00
17	(Q47A) Essential to Have House Warm	17		0.02	0.00
18	(Q25A) Home Unoccupied For More Than 7 Days	18		0.02	0.00
19	(Q2C1) Added Wall Insulation		17	0.02	0.00
20	(Q3E1) Number of Water Heaters	18		0.01	0.00
21	(Q42A) Broken Equipment	19		0.01	0.00
22	(Q2C1) Added Wall Insulation	20		0.003	0.00
23	(Q17) Bedroom Window Kept Open in Winter		19	0.003	0.00
24	(Q47E) Willing to Wear Heavy Clothing		18	0.003	0.00
25	(Q2E1) Added Floor Insulation	19		0.002	0.00
26	(Q47B) Want to Wear Light Clothing	20		0.001	0.00

A3c. Canonical Discriminant Function

Eigenvalue	Canonical Correlation	Wilks' Lambda	Chi-Squared	Degrees of Freedom	Significance
899.91	0.99	0.001	108.85	20	0.00

A3d. Correlations Between Discriminating Variables and Canonical Discriminant Function

Q2B1	0.02
Basement	-0.01
Q3J1	0.01
Q47B	0.01
Q2C1	0.01
Q47A	0.01
Q2G1	-0.01
Q25A	-0.01
Q48	0.01
Montana	-0.01
Q3F1	-0.01
Q2E1	0.01
Q3E1	-0.01
Size of Household	-0.004
Q3I1	0.004
Space	-0.004
Q53	-0.003
Q42A	0.002
Q47F	0.002
Q1A	0.0001

A3e. Canonical Discriminant Function Coefficients

	Unstandardized	Standardized
Q1A	3.27	7.47
Q2B1	42.80	16.48
Q2C1	-12.99	-4.55
Q2E1	5.36	1.70
Q2G1	-44.77	-11.62
Q3E1	49.51	9.40
Q3F1	-28.37	-5.39
Q3I1	16.68	29.03
Q3J1	7.00	13.06
Q25A	-5.29	-5.08
Q42A	-13.68	-5.81
Q47A	3.89	3.30
Q48B	-1.71	-1.83
Q47F	-14.38	-11.60
Q48	32.98	30.46
Q53	-6.96	-5.72
Size of Household	16.68	17.72
Montana	28.99	5.50
Basement	-48.60	-17.67
Space	-0.003	-17.71

A3f. Classification Results

Actual Group	Number of Cases	Predicted Group Membership H	L
Group H	16	16	0
		100.0%	0.0%
Group L	38	0	38
		0.0%	100.0%

Percent of Grouped Cases Correctly Classified: 100.0%

Reduced models

The model for all households was composed of 30 variables and was considered unwieldy. Accordingly, we decided to test a reduced form of the model by concentrating on a subset of the variables that would give us a satisfactory model. We ran ten models using subsets of variables taken from the complete analysis (Table A4 shows the results of the first four models). The basis for including the variables in a subset was the value of the Wilks' lambda: only the highest ones were included. The first model included only one variable, the second model included two variables, the third three variables, and so on.

Table A4. Reduced Models for All Households

Variable	Model Number			
	1	2	3	4
Installed weatherstripping	x	x	x	x
Winter outdoor temperature		x	x	x
Number of electric space heaters			x	x
Presence of a central thermostat				x
Percent of Group H Correctly Classified	20.8	79.2	79.2	83.3

The measure of success for each model was its ability to successfully classify the high outliers. Table 4 shows that the floor area model was able to classify 60% of the high outliers correctly. Models with five or more variables did not improve our ability to correctly classify Group H households. The model with four variables was able to classify 83% of the Group H households. Since the complete model was able to classify 100% of the cases, the model with four variables was less accurate, however, still valuable in classifying Group H households.