Controlling for Program Participation Self-Selection Bias

Shawn Bodmann, DNV KEMA Energy & Sustainability, Madison, WI
Co-authors:
Tammy Kuiken, DNV KEMA Energy & Sustainability, Madison, WI
William Golemboski, DNV KEMA Energy & Sustainability, Burlington, MA
Mitch Rosenberg, DNV KEMA Energy & Sustainability, Burlington, MA
Miriam Goldberg, DNV KEMA Energy & Sustainability, Madison, WI
Robin Sempf, Alliant Energy

ABSTRACT

This paper describes an evaluation of the energy impacts of an informational program available to residential customers. The program is a utility-produced home improvement show (PowerHouse™) that promotes energy efficiency. The evaluation used a combination of surveys and other cost-effective information to estimate savings attributable to the PowerHouse show, and attempted to minimize self-selection bias by controlling for possible a priori attitudinal and demographic differences between viewers and non-viewers.

The evaluation produced evidence that viewers were different from non-viewers on characteristics that affected energy use and energy savings. This finding supported the decision to control for these differences in the regression models. In particular, the evaluation estimated a propensity to be a viewer, so that effects of program viewing on energy saving behaviors could be determined controlling for this pre-disposition.

The regression models resulted in estimated savings attributable to the PowerHouse show of 10 kWh annually and 40 watts per viewing household, after controlling for differences in viewing propensity and other characteristics. All attributable savings occurred from end uses not covered by existing "downstream" rebate or audit programs, mostly lighting. The total estimated kWh savings were about the same magnitude as the insulation and infiltration portion of Alliant Energy's 2009 home energy audit program, and the kW savings were almost equivalent to their 2009 residential prescriptive rebate program.

Several sources of uncertainty remain, including: the presence of an "upstream" lighting rebate program, the underlying structure of the dependent variables, and a reliance on self-report data for certain key variables.

Introduction

The evaluation industry has long struggled with the practice of using energy efficiency program[1] non-participants as a proxy for what would happen absent such programs. People who choose not to participate in energy efficiency programs (non-participants) are not necessarily an effective comparison group because of the possibility that underlying pro-efficiency attitudes and demographics lead people both to participate in energy efficiency programs and to undertake efficiency projects. These underlying differences can inflate energy efficiency program effects estimated via simple participant versus non-

[1] This paper uses the word "program" to refer to energy efficiency programs, and the word "show" to refer to the PowerHouse television show.

participant comparisons. Evaluation techniques relying instead on market-level data or regional comparisons may be less vulnerable to these biases, but can also be expensive and are not always practical. Statistical techniques to control for possible self-selection effects that can be applied to survey-based studies can provide a cost-effective way to improve the validity and accuracy of energy efficiency program evaluations based on participant/nonparticipant comparisons.

PowerHouse Show

For 16 years, Interstate Power and Light Company (IPL), an Alliant Energy company,[2] has produced PowerHouse™, an educational television show that focuses on energy efficiency for the home. In addition to the television show, PowerHouse has a website that provides supplemental articles, energy-related facts, energy savings calculators, and links to other informational websites such as those with information on Alliant Energy rebates. PowerHouse provides strong brand identification for Alliant Energy and a vehicle for customer outreach.

The purpose of PowerHouse is to educate customers on energy efficiency and encourage implementation of energy efficiency measures. The show offers useful tips on reducing home energy use through more efficient heating, cooling, lighting, and more. In addition to airing once per week across four Iowa television markets, show segments are available on the PowerHouse website and the Alliant Energy PowerHouse TV channel on YouTube. The PowerHouse hosts are also involved in community promotions such as CFL campaigns, local farmers' markets, and a breast cancer awareness campaign. PowerHouse is advertised through bill inserts, all residential promotional materials (such as a residential rebate booklet), radio, web banners, and some newspaper ads. PowerHouse DVDs are also sent to schools and libraries in Alliant Energy service territories.

PowerHouse Evaluation Overview

Alliant Energy hired the evaluation team to estimate the energy savings attributable to its PowerHouse television show's influence on viewers' behaviors and energy saving actions, both those that were already incentivized by existing Alliant Energy rebate and audit programs and those that were not. An important part of this evaluation was to quantify the extent and effects of self-selection in PowerHouse viewing. This step was necessary to address the possibility that attitudinal, demographic, or other factors that increased the tendency to be a PowerHouse viewer would also increase the tendency to save energy even without PowerHouse. To avoid over-estimating the energy efficiency behaviors attributable to PowerHouse, the evaluation team needed to estimate and control for the PowerHouse self-selection effects.

The primary customer-specific data used for the analysis consisted of:

- Tracking data on participation in rebate and audit programs,
- Customer surveys identifying viewership, energy saving actions taken, and demographic and attitudinal characteristics, and
- Nielsen market segmentation data that provided additional demographic information

The key analysis had three steps. First (1) the evaluation team estimated energy savings for the energy saving actions reported on the customer surveys. Next (2) the evaluation team used binary logistic regression to estimate the effect of various demographic and attitude variables on PowerHouse

[2] Alliant Energy operates in Iowa, Wisconsin, and Minnesota, and the PowerHouse show airs in six television markets across those three states. However, the evaluation described in this paper focused only on Iowa customers and viewers.

viewing; this regression produced a variable estimating the propensity to view PowerHouse. Third (3) the evaluation team estimated end-use energy savings attributable to PowerHouse viewing while controlling for propensity to be a viewer by regressing the end-use energy savings (from step 1) on a set of variables that included the propensity-to-view variable. Application of the savings regression estimates (from step 3) to the Alliant Energy customer base produced estimates of per-viewer and total population savings attributable to PowerHouse.

Survey Design

The evaluation's primary data source was telephone surveys where respondents (600 viewers and 605 non-viewers) reported PowerHouse viewing, over 50 different recent energy saving actions, energy efficiency attitudes (EEA), and demographics. Additional data sources included Alliant Energy rebate and audit program participation data and market segmentation data from Nielsen.

PowerHouse Viewing

The survey included two questions to determine whether a respondent was a PowerHouse viewer. Respondents who answered yes to either of these questions were classified as viewers.[3]

- Have you watched Alliant Energy's PowerHouse television show anytime in the past two years?
- Have you watched segments of Alliant Energy's PowerHouse TV show on the Internet such as YouTube or the PowerHouse website anytime in the past two years?

Energy Efficiency Attitudes

The survey included a measure of attitudes and knowledge about energy efficiency (the EEA scale) to help identify the people that might be predisposed to view PowerHouse and to perform energy efficiency behaviors. The survey included three attitude questions measured on a five-point scale anchored at "strongly disagree" and "strongly agree", and with a "neutral" option in the middle of the scale:

- Conserving energy is important to me.
- I want to reduce my household energy use to lower my utility bill.
- I want to reduce my household energy use to protect the environment.

The survey also included three knowledge questions:

- Before today, had you heard of ENERGY STAR?
- To the best of your knowledge, which of the following activities saves the MOST energy?
- Which one saves the LEAST energy?
  - Replacing old appliances like refrigerators
  - Installing more or better insulation in your home
  - Turning off the lights when you leave the room

The survey provided respondents with the same three options for each of the last two knowledge questions. Respondents could receive up to five points in the energy knowledge scale: one point for answering yes to the ENERGY STAR question, two points for correctly identifying the action that saves the most energy (insulation), and two points for correctly identifying the activity that saves the least

[3] Nielsen ratings could not be used to determine viewership because they were not available at a specific household level.

(turn off lights). The sum of this knowledge score (0 to 5) and the average of the three attitude questions (1 to 5) provided a composite Energy Efficiency Attitudes (EEA) score (1 to 10).

Energy-saving Behaviors

Respondents answered whether they had performed any of a large number of energy saving actions in the past two years. To make the questioning manageable, the surveys first asked the respondents whether they had done any actions within a certain category, and only upon an affirmative answer asked about specific end uses within each category. The survey included questions for the following categories:

- Appliances
- Building Shell
- Electronics
- HVAC
- Lighting
- Water Heating
- Other

Demographics

The final type of information collected through the surveys was demographics, including:

- Confirmation of Alliant Energy service type (electric, gas, both)[4]
- Own or rent home
- Education of respondent
- Household income

A purchased battery of Nielsen market segmentation data for each survey respondent included income, financial donations to environmental causes, and do-it-yourself hobbies. The Nielsen data had certain limitations. Income data for about half of the sample were reported, and for the other half were imputed by Nielsen. The donations and do-it-yourself variables were reported, but missing for the majority of our sample. The analysis treated the missing values as zeroes.

Key Analyses

The evaluation team followed a three step analysis process. The first step was to calculate energy savings for each respondent. The second step was to estimate a propensity to view PowerHouse for each respondent to represent self-selection effects (the PHV Correction Factor). The final step was to estimate the amount of the savings from step 1 that was attributable to PowerHouse viewing, after controlling for the self-selection effects calculated in step 2 (Figure 1).

[4] The evaluation team had these data from the Alliant Energy billing records, and we also confirmed the data with the surveys.

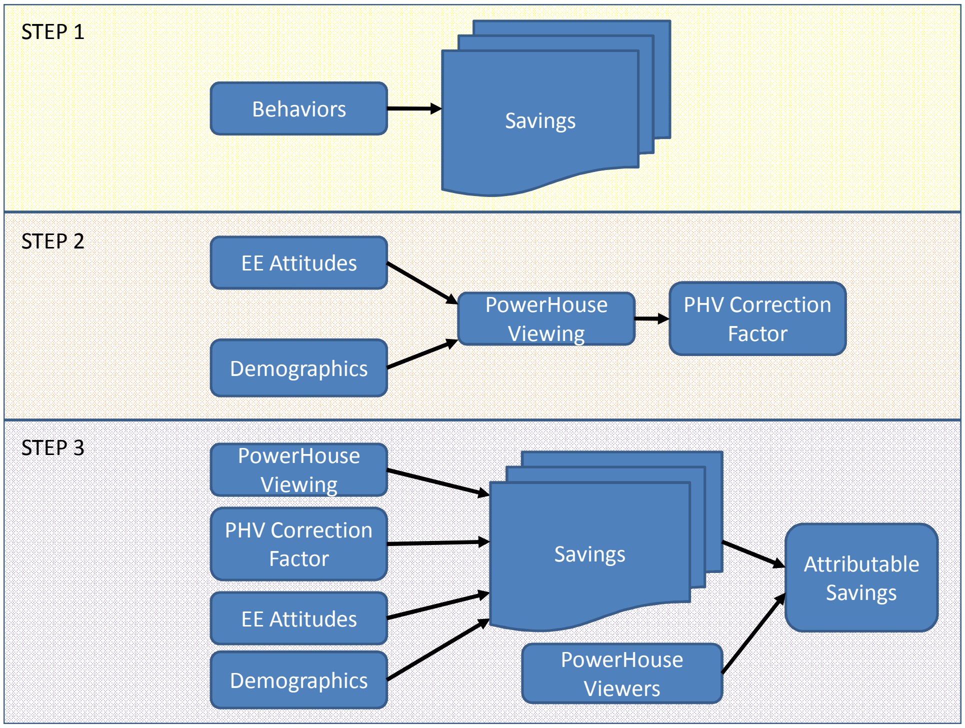

Figure 1. Analysis Steps

Step 1 – Calculate Savings

Engineers on the evaluation team calculated a typical savings value (kWh and kW, or therms) for each specific end-use action, and then bundled those savings into a total for each end-use category. In cases where Alliant Energy had an audit or rebate program with a deemed savings estimate, the engineers used the deemed savings estimates. The engineers used Alliant Energy records to divide savings into those that occurred outside existing audit or rebate programs (untracked savings) and those that were already tracked by an existing audit or rebate program (tracked savings). The engineers then summed untracked and tracked savings for each end-use category for each respondent by energy unit. This resulted in a 42 cell matrix of savings estimates for each respondent (Table 1). Checks indicate the 37 combinations that the evaluation team estimated savings for.

Table 1: Energy Savings Categories

End-Use Category	Untracked Savings			Tracked Savings		
	kWh	kW	Therms	kWh	kW	Therms
Appliances	✓	✓	✓	✓	✓	✓
Building Shell	✓	✓	✓	✓	✓	✓
Electronics	✓	✓	✓			
HVAC	✓	✓	✓	✓	✓	✓
Lighting	✓	✓		✓	✓	
Water Heating	✓	✓	✓	✓	✓	✓
Other	✓	✓	✓	✓	✓	✓

Step 2 – Calculate PowerHouse Viewing Propensity and Correction Factor

The second analysis step was to calculate a PowerHouse Viewing Correction Factor at the household level, to use later to account for any PowerHouse viewing self-selection effects. This step used a binary logistic regression to determine a customer's propensity for viewing PowerHouse, based on several demographic variables (Equation 1).

Equation 1: Model Predicting PowerHouse Viewing

$$PH_i = b_0 + b_1 * attitudes_i + b_2 * income_i + b_3 * homeowner_i + b_4 * college_i + b_5 * BTU_i + b_6 * greendonor_i + b_7 * DIYhobby_i + b_8 * electric_i + b_9 * gas_i$$

where:

i =	Subscript for each individual household.
PH =	0 if household did not view PowerHouse in last 2 years, 1 if household viewed at least one segment of PowerHouse in last two years. This variable was based on survey responses.
attitudes =	Scale combining EEAs and knowledge questions from the surveys. Each respondent could score a 0-10 on the scale, which the evaluation team converted to deciles and included the deciles as a rank-ordered categorical variable in the model.
income =	Household income converted to four-level rank ordered categorical. Taken from survey response if response was given, else obtained from Nielsen data.
homeowner =	0 if respondent did not own their home, 1 if respondent did own their home. This variable was based on survey responses.
college =	0 if respondent had no college education, 1 if respondent had some college. This variable was based on survey responses.
BTU =	Total billed household energy usage based on billing data, combining electric and gas usage, in BTU, then converted to a rank-ordered categorical variable by decile. Energy usage covered the same time period for all respondents and was not corrected for climate differences.
greendonor =	0 if household did not make donation to a pro-environmental or pro-wildlife cause, 1 if household did make such a donation, from Nielsen.
DIYhobby =	0 if no household members reported having a do-it-yourself hobby, 1 if someone in household did report such a hobby, from Nielsen.
electric =	0 if customer did not purchase electricity from Alliant Energy, 1 if customer purchased electricity from Alliant Energy. This variable was based on billing data.

gas = 0 if customer did not purchase gas from Alliant Energy, 1 if customer purchased gas from Alliant Energy. This variable was based on billing data.

The output of the first model provided the PowerHouse Viewing Correction Factor (C) value for each customer household (i). This variable helped to isolate the effect of PowerHouse from the respondents' characteristics that led to both PowerHouse viewing and energy saving actions. The viewer correction factor is described by the following equation:

Equation 2: Viewer Correction Factor Calculation

$$C_i = \frac{[\hat{P}_i * \ln \hat{P}_i]}{(1 - \hat{P}_i)} + \ln \hat{P}_i$$

where:

i	=	Subscript for each individual household
$\hat{P}_i$	=	The fitted value for the dependent variable (i.e. binary variable for PowerHouse viewing) in the Logit of the PowerHouse viewership model.

The basic idea was to purge the PowerHouse viewer status variable in Equation 1 of the effects of self-selection by estimating a participation model for PowerHouse viewing and applying this instrumental variable back into the energy savings models (in Step 3).

This adjustment technique was first introduced in Heckman (1978), and refined and expanded in Dubin and McFadden (1984). Self-selection correction methods have been applied by various authors in the context of energy efficiency savings evaluations (Train, et al. 1994; Goldberg & Kademan 1995; Goldberg, Michelman, & Dickerson 1997). It is generally recognized by practitioners (e.g.: XENERGY 1995) that correction factors such as the one used here are imperfect: they depend on distribution assumptions that cannot be validated and are not mitigated by large sample sizes, and in some cases may be worse than no correction. At the same time, any relatively simple regression model that describes behavior and energy consumption in terms of physical and demographic characteristics is only an approximation to a complex set of processes. The models used here represent a reasonable and useful attempt to control for observable factors that tend to be different between viewers and non-viewers, including their underlying propensity to be viewers. The model results for Equation 1 indicate that certain household characteristics are indeed associated with higher or lower tendency to be a viewer. The statistical significance of the coefficients on the correction factor term C in Equation 3 below indicates that these terms are capturing an effect not fully accounted for by the other predictor variables alone in linear form. While a variety of refinements could be considered for this analysis, it yields results as described below that appear to be meaningful.

Step 3 – Calculate Savings Attributable to PowerHouse

After developing the correction factor (C), the evaluation team estimated a set of ordinary least squares (OLS) regression models for each combination of 37 energy type (kW, kWh, therm) by end-use category (Appliances, Building Shell, Electronics, HVAC, Lighting, Water Heating, Other) by tracked/untracked status listed in Table 1. For each regression fit, the dependent variable in Equation 3 was the savings for that category. These models predicted the influence of PowerHouse viewing on the energy savings resulting from each end use, controlling for propensity to be a viewer.

Equation 3: Model Predicting Energy Savings

$$Savings_{i_j} = b_0 + b_1 * attitudes_i + b_2 * income_i + b_3 * homeowner_i + b_4 * college_i + b_5 * BTU_i + b_6 * greendonor_i + b_7 * DIYhobby_i + b_8 * electric_i + b_9 * gas_i + b_{10} * C_i$$

Equation 3 regressed each energy unit by end-use by tracked/untracked savings (j) for each household (i) on all of the variables from Model 1 and the household level viewer correction factor (C_i). Constructing the models like this had the following advantages:

- Bundling into end-use categories increased the potential for producing statistically significant results by increasing the number of non-zero savings values.
- Bundling also created greater internal consistency on the results by combining similar energy efficiency actions together. For example, the lighting category included: turning off lights more often, using occupancy sensors, and using lighting timers.
- Modeling each energy unit (kWh, kW, therms) separately allowed the attributable savings to be reported separately.
- Including the per-unit energy savings in the model produced output that directly reports the marginal realized savings, rather than measure adoption probabilities.
- Including the correction factor C reduces the potential for self-selection bias in the resulting attributable energy savings estimates.
- The model results could be applied to a broader population of residential customers served by Alliant Energy.

Finally, the output of the OLS regressions provided population-level estimates of the effects of viewing PowerHouse, using the following rules:

- For end-use categories for which the OLS models showed that PowerHouse viewing had a statistically significant effect on energy savings (at the 90% confidence level), multiplying the savings effect of PowerHouse viewing (PowerHouse viewing coefficient from the OLS regression) by the estimated number of PowerHouse viewers in Alliant Energy service territory (about 125,000) gave the total savings attributable to PowerHouse viewing.
- For end uses where the PowerHouse viewing effect from the OLS regression was not statistically significantly different from zero, the evaluation team used an estimate of 0.

Key Results

PowerHouse Viewing

About one-fourth (28%) of the survey respondents reported viewing the PowerHouse show at least once in the last two years. Multiplying this percentage by the total number of customers within the studied viewing area produced an estimate of about 124 thousand viewing households. Of the viewers, about half reported viewing the show less than once per month and about 40 percent reported viewing the show at least once per month.

Energy Efficiency Attitudes and Actions

EEA scores correlated both with taking energy efficiency actions and viewing PowerHouse. Respondents who took at least one energy efficiency related action had higher average EEA scores (6.7) than those who did not take energy efficiency actions (5.8). PowerHouse viewers had a higher EEA score (6.8) than non-viewers (6.2). Viewers were also statistically significantly (at the 90 percent confidence level or higher) more likely than non-viewers to take 30 out of the 54 surveyed energy saving actions.

The combination of these differences supported the hypothesis that PowerHouse viewers were subject to influences that both increased their propensity to view PowerHouse and increased their propensity to do energy efficiency. This result indicates that non-viewers were not the best representation of what viewers would do in the absence of PowerHouse, and a strict difference in installed savings between viewers and non-viewers would not necessarily produce the energy savings attributable to the PowerHouse show.

Regression Results

The estimates of population level savings attributable to PowerHouse viewing are shown in Table 2. The electronics and lighting end-use categories had statistically significant kWh savings untracked by existing rebate or audit programs. The building shell, electronics, HVAC, and lighting categories had statistically significant untracked kW savings. Dashes represent categories that did not have statistically significant results. There were no statistically significant effects for untracked therms savings.

Per household savings were about 10 kWh annually and about 40 watts. This magnitude of savings is about the same as replacing one incandescent lamp with one CFL and seems reasonable for a “treatment” as minor as viewing a television show at least once in the last two years.

Summing across the PowerHouse viewing population of about 124 thousand households resulted in an estimated total annual savings of approximately 1.2 million kWh and 4.7 thousand kW. The estimated total kWh savings were roughly equivalent to the insulation and infiltration portion of Alliant Energy’s 2009 home energy audit program while the kW savings were almost equivalent to their 2009 residential prescriptive rebate program. Thus, while the per-household results were modest, the ubiquity of the program resulted in material overall savings.

While it is possible that the PowerHouse show could increase participation in existing Alliant Energy audit and rebate programs, we did not find any statistically significant savings associated with end uses that were already tracked by existing programs. While this result may be disappointing, it does mean that all of the savings we reported were from untracked sources.

Table 2: Savings Attributable to PowerHouse by End Use

End Use Category	Untracked Savings	
	kWh	kW
Appliances	-	-
Building Shell	-	5,474
Electronics	457,802	29
HVAC	-	3,800
Lighting	2,025,880	88
Water Heating	-	-
Other	-	-
Total	2,483,682	9,390

Key Uncertainties

There were a number of key uncertainties regarding the results in Table 2.

- *Estimated attributable lighting savings may result in part from existing upstream energy efficiency programs.* Alliant Energy participates in a statewide upstream lighting program that reduces the cost of CFLs at the point of sale but does not consistently track customer participation in a way that allowed the evaluation team to identify upstream lighting program participants as part of this study. It is likely that some of the untracked lighting savings reported in Table 2 were influenced by and already claimed by the upstream program, but we had no way of estimating how much. Regardless of the amount of this overlap, this evaluation still produced evidence that the PowerHouse show affected these purchases. The estimated savings attributable to PowerHouse were a small portion (about 4%) of those claimed by the upstream program.
- *The structure of the dependent variables (savings) were not ideally described by the statistical models the evaluation team used.*
  - The dependent variables were nonnegative and have large numbers of zeros. This structure is not ideally described by a linear model with symmetric errors such as the one the evaluation team used. Possible alternatives would be a form of Tobit bounded below, or a multinomial interacted with a continuous term. These more complex approaches were beyond the scope of this study.
  - For many of the measures studied, the savings variable had several distinct levels, with a moderate number of small nonzero values and a small number of much larger nonzero values. Data of this structure in particular is not well suited to the linear model with symmetric errors.
- *The self-selection term in this case can be thought of as a proxy for propensity to be a viewer, but can't be assumed to completely eliminate the self-selection problem.* The PowerHouse viewing correction factor is approximately correct if the residual errors are normally distributed. With large numbers of zeros in the original data and lumpy non-zero values, the residual errors were not normally distributed.
- *The models of savings that occurred within tracked rebate or audit program did not produce statistically significant results.* The tracking data tended to have very few nonzero cases for either viewers or non-viewers, so we could not find statistically significant savings. Our

evaluation failed to produce evidence that PowerHouse affected participation in these programs, but it might have.

- *Some of the key variables were more reliable than others.* More reliable variables included the rebate tracking data and self-reported basic household characteristics. Recall of PowerHouse viewing and of energy saving actions performed over the past two years are probably less reliable. The Nielsen data provided reported values when available, and imputed or missing values otherwise. This analysis treated the missing values as zeroes, which can bias results.
- *Rigorous validation of some of the composite scales used in the regressions models was not practical.* This included the key EEA scale.
- *The engineering estimates of savings required several assumptions and averaging.* The evaluation team used deemed savings from existing Alliant rebate or audit programs whenever possible. However, the nature of the data collection made it impossible to get enough information to compute detailed savings estimates for every specific end use.

Despite these uncertainties in the savings results, the study provided qualitatively useful findings, and an approach that can be applied and refined in other evaluations of open-access information programs. The study produced evidence that people who choose to watch the PowerHouse show were apt to also engage in energy saving actions. However, after partially controlling for those potential sources of bias, there remained evidence that the PowerHouse show contributed to greater levels of energy savings than would happen in the absence of the television show and associated internet site. The attributable savings, while modest on an individual household level, add up to a similar magnitude as other more traditional rebate or audit programs funded by Alliant Energy. This evaluation also provides a case-in-point example of a technique to partially control for self-selection biases that may factor into evaluations of other energy efficiency programs.

References

Dubin, J. & McFadden, D. 1984. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52 (2): 345-362.

Goldberg, M., and E. Kademan, 1995. "Is It Net or Not? A Simulation Study of Two Methods." *In Energy Program Evaluation: Uses, Methods, and Results,* 459-465. Chicago, IL: National Energy Program Evaluation Conference.

Goldberg, M., T. Michelman, C.A. Dickerson, 1997. "Can We Rely on Self-Control?" *In The Future of Energy Markets, Evaluation in a Changing Enviornment,* 187-197. Chicago, IL: National Energy Program Evaluation Conference, August 1997.

Heckman, J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica* 46 (4): 931-959.

Train, K., S. Buller, B. Mast, K. Parikh, E. Paquette, 1994. "Estimation of Net Savings for Rebate Programs: A Three-Option Nested Logit Approach." *In Proceedings from the ACEEE 1994 Summer Study on Energy Efficiency in Buildings,* 7:7.239-7.247. American Council for an Energy-Efficient Economy, Washington DC.

XENERGY, Inc. 1995. Net Savings, Estimation: An Analysis of Regression and Discrete Choice Approaches. Paper prepared for CADMAC Subcommittee on Base Efficiency.