NEW JERSEY CONSERVATION ANALYSIS PROJECT: ENERGY SAVINGS FROM RESIDENTIAL PROGRAMS

Michael Ozog, Daniel Violette, and J. Sueker
RCG/Hagler, Bailly, Inc.
Boulder, Colorado

Introduction

The purpose of this paper is to discuss the method used to analyze a series of utility conservation programs mandated by the New Jersey State Legislature. The project is being conducted for the New Jersey Conservation Analysis Team (NJCAT).[1] NJCAT is a consortium of the seven investor-owned gas and electric utilities in the state, the Board of Public Utilities, the Public Advocate, and the Energy Division of the Department of Commerce.

This project is one of the largest conservation program evaluations undertaken, encompassing an integrated investigation spanning several utilities and several programs. Data collection alone involved several different firms working for more than one year. This project is noteworthy not only for the magnitude of the evaluation, but also for the innovative approaches used to estimate energy savings. These estimates are obtained by both statistical models and engineering models. The statistical modeling involves both discrete choice modeling and multivariate regressions, and is designed to address significant evaluation problems such as how to control for self-selection bias. The engineering simulation estimates energy savings by using building energy-use analysis models. Combining the statistical model with an engineering simulation allows estimation of both total energy saved and reductions in peak demand.

This paper will concentrate on methodology. The next section gives a brief description of the various conservation programs in this project. The discussion continues with an outline of the methodology used to evaluate these programs, addressing data collection, quantitative formulation, and concluding with an outline of the engineering simulation.

Scope of the NJCAT Project

Tables 1 and 2 list the programs being evaluated for each of the seven utilities. Though there are 54 separate programs, these can be aggregated into four distinct types: the Home Energy Savings Program (HESP), the Seal-Up program, an Appliance Rebate program, and the Low Income Direct Grants Program.

The HESP program is a residential program which involves an in-house energy audit by a New Jersey certified auditor. The auditor supplies the house with a report identifying conservation measures, their costs, and the expected savings. If the customer decides to undertake any of these recommended measures, they have the option of applying for the Low/No Interest Loan program. This program allows customers to receive loans for the conservation measures at no interest if the household's income is less than $30,000, or a subsidized interest rate which varies by utility if the family earns more than $30,000. Since an audit must be conducted before the customer is eligible for the subsidized loan, these two programs are pooled together in this analysis.

The Seal-Up program involves a slight fee of approximately $10, for which the utility wraps the customer's electric water heater and turns down the water heater thermostat. If this has already been done, the fee can be used for other water heater-related actions or weatherstripping actions. "Assistance" Seal-Ups are seal-up programs provided free to eligible low-income or senior citizen customers.

The Appliance Rebate program provides incentives to residential customers who install or replace energy-efficient appliances. The type of appliance covered differs depending on the type of program adopted by each utility. For example, Atlantic Electric has a rebate program for air conditioning, where a customer receives a rebate for purchasing an air conditioning unit with high EER/SEER. The amount of the incentive is proportional to the EER/SEER of the new unit.

The Low Income Direct Grants Program provides direct weatherization assistance to low-income, electrically heated residential customers. The utility pays for the completion of conservation activities up to a predefined limit, which varies from utility to utility.

There are several advantages to conducting a multiple program analysis. First, one control group can be used for all of the programs, rather than one for each program. Second, there are efficiencies gained in questionnaire administration since there are many basic customer characteristics questions which are common to each program. This saves on sampling and questionnaire design. Third, the large amount of data available allows for an improve-

Table 1. Electric Utility Conservation Programs to be Evaluated

Atlantic Electric	Jersey Central P&L	Rockland	PSE&G Electric
Home Energy Savings	Home Energy Savings	Home Energy Savings	Home Energy Savings
Low/No Interest Loan	Low/No Interest Loan	Customer Seal-up	Low/No Interest Loan
Conservation Seal-up	Wrap-up/Seal-up	Low-income Seal-up	Customer Seal-up
Low-income Seal-up	Low-income Seal-up	Senior Save	Senior Save
Senior Save	Senior Save	Air-conditioning Rebate	Winter Termination
Winter Termination	Winter Termination	CACS	Air-conditioning Rebate
Air-conditioning Rebate	Air-conditioning Rebate		Heat Pump Rebate
Heat Pump Rebate	CACS		CACS
CACS	Commercial/Industrial Customer Contact		Low-income Direct Grants
Low-income Direct Grants	Commercial/Industrial A/C Rebate		
	Commercial/Industrial Lighting		
	Commercial/Industrial Fluorescent Ballast		
	Low-income Direct Grants		

Table 2. Gas Utility Conservation Programs to be Evaluated

PSE&G Gas	Elizabethtown Gas Co.	New Jersey Natural	South Jersey Gas
Home Energy Savings	Home Energy Savings	Home Energy Savings	Home Energy Savings
Low/No Interest Loan	Low/No Interest Loan	Low/No Interest Loan	Low/No Interest Loan
Customer Seal-up	Conservation Seal-up	Conservation Seal-up	Residential Seal-up
Low-income Seal-up	Low-income Seal-up	Low-income Seal-up	Low-income Seal-up
Senior Save	Senior Save	Senior Save	Senior Save
Winter Termination	Winter Termination	Winter Termination	Winter Termination
Gas Heater Rebate	Gas Furnace Rebate	Gas Furnace Rebate	Gas Heater Rebate
Thermostat Rebate	Water Heater Rebate	Water Heater Rebate	Water Heater Rebate
CACS	Thermostat Rebate	Thermostat Rebate	CACS
Low-income Direct Grants	CACS	CACS	Low-income Direct Grants
	Low-income Direct Grants	Low-income Direct Grants	

ment in the precision of energy savings estimates. Finally, a multiple program analysis allows comparison between programs to see which is more effective in reducing energy consumption.

Data Collection

The first step in this analysis was to obtain the billing histories for an initial sample of participants. This was done to assure complete histories would be available for every survey respondent.

One issue in data collection is the length of time for which billing data is required. The issue is how much historical data on energy use is required to estimate properly post-program energy savings. At a minimum, 12 months of billing data prior to program participation and 12 months of post-participation billing data are needed. However, it is desirable to have several post-participation time periods since there is some concern regarding the magnitude of program-induced energy savings in subsequent years. Therefore, billing data was collected for three years: one year prior to participation, one year after participation, and two years after participation.

Choosing the control group was done by requesting the billing records of the customers just before and just after each participant, checking to see if those customers were also participants. Since billing data are generally organized on the basis of meter-reading routes, this technique was done to obtain a control group of nonparticipants who matched participants. One problem that occurred with this is that some of the low-income programs involved weatherizing an entire housing project or neighborhood, which would generally leave few non-participants. Therefore, to get enough low-income and senior citizen customers, the control sample was augmented with a random sample of "assistance-eligible"[2] customers.

Once the billing data were collected, it was screened for those customers with an insufficient amount of billing data, or who had one or more billing entries which appeared far out of line. Once these customers were eliminated, a survey sample was selected from the remaining accounts. Based upon estimates of the necessary sample size for estimation efficiency and hypothesis testing, it was decided that the sample size should involve over 14,000 customers, so the survey size was set accordingly. Table 2 shows the count of completed

Table 2. Completed Questionnaires

Program	Atlantic Electric	Jersey Central P&L	Rock-land	PSE&G Electric	PSE&G Gas	Elizabeth town Gas Col	New Jersey Natural	South Jersey Gas	TOTAL
Control	514	735	78	197	772	288	228	231	3043
Control, Low-income	178	225	0	57	286	20	88	92	946
Total Control	692	960	78	254	1058	308	316	323	3989
HESP	304	553	62	180	680	265	116	295	2455
HESP Loan	0	0	0	3	133	33	20	56	245
Total HESP/Loan	304	553	62	183	813	298	136	351	2700
Seal-up, Unspecified	0	385	0	88	524	0	0	147	1144
Seal-up, Regular	267	0	0	88	315	97	69	22	858
Seal-up, Assistance	92	0	0	0	0	216	76	0	384
Seal-up, Supplements	0	1512	0	46	129	54	10	14	404
Total, Seal-up	359	536	0	222	968	367	155	183	2790
Rebate, Unspecified	0	0	0	0	0	0	227	106	333
Rebate, Space Heat	28	66	0	61	126	78	0	0	359
Rebate, Air Conditioner	163	304	49	0	327	0	0	0	843
Rebate, Water Heater	0	0	0	2	103	108	0	0	213
Rebate, Thermostat	0	0	0	14	97	89	101	0	301
Rebate, Supplements	0	9	0	4	133	45	11	49	251
Total Rebate	191	379	49	81	786	320	339	155	2300
TOTAL ALL PROGRAMS	1546	2438	189	740	3625	1293	946	1012	11779

questionnaires by program. Overall, the survey achieved a 72% response rate.

The design and implementation of the survey portion of the analysis was undertaken with WESTAT, a research corporation in Maryland. During this phase of the project, several interesting issues arose. First, the New Jersey State Legislature, in mandating these programs, also required electric utilities to provide audits and seal-ups to customers that had any type of space heating other than gas. This meant that individuals who use oil for space heating may also be participants in the conservation programs. However, the statistical analysis requires data on actual energy use, which is not readily available for users of oil space heating. Therefore, it was decided that for the HESP, Low Interest Loans, and Seal-Up programs, questionnaires would only be sent to those customers who had space heating data. Oil savings associated with the audit/loan and seal-up programs are estimated using a proxy estimate obtained from the gas utility estimates.[3] For the appliance rebate program, the fuel saved is electricity or gas independent of the type of fuel used for space heating, so a questionnaire was sent regardless of the space heating fuel used by the customer.

The other issue involved sampling from Public Service Electric and Gas Company (PSE&G). Since this utility is both an electric and gas utility, it was decided to over-sample customers that had both electricity and gas services provided by PG&E. This was done to obtain a richer energy use data set.

The questionnaires were designed to accomplish three things:

1. Gather data on important factors that affect energy use, such as household size and income.
2. Determine how these factors changed over the three time periods of estimation.
3. Capture data on the customer's attitudes, opinions and beliefs which might affect the probability of participation and taking conservation actions. This also included questions regarding the customer's expectations about future energy prices as well as the payback requirements.

The latter two features are relatively unique to this particular study. The questions regarding changes over time in the factors that affect energy use were designed to aid in the estimation of the change in energy consumption over the three time periods. The incorporation of questions about factors which might affect participation was designed to aid in the probabilistic estimation of program participation, which is used to correct for self-selection bias. These issues are explained in the next section.

Another data set available for analysis was a computerized version of the HESP auditor's report. This report included a list of recommended actions, the auditor's cost of each action, and the auditor's estimated savings from taking the action. In addition, the auditor's estimates of energy savings can be compared with the statistically estimated energy savings.

One final set of data was daily temperature for use in weather normalizing monthly energy consumption. This is necessary for the Princeton Scorekeeping Method (PRISM). PRISM uses normalized annual energy consumption (NAC) as a consumption index from which energy savings and conservation trends may be accurately estimated. In this project, PRISM was used both as a comparison method, where the NAC for participants both before and after the program is compared with the NAC for non-participants over the same periods, and as a pre-processor for the statistical analysis. As a pre-processor, it was used to normalize for weather during the heating/cooling seasons.

Quantitative Formulation

A fundamental problem associated with estimating the effect of any conservation program involves the measurement of the energy consumed by participants before and after participation to determine the change in consumption brought about by the program. One difficult issue which must be resolved is how much energy would have been consumed by the participating households had the program not been in effect. It is likely that some households would have taken conservation measures even if the utility's conservation program had not been in place. This "natural" conservation should not be inadvertently attributed to the program.

If, as is the case in this project, observations are available on both program participants and non-participants before and after program implementation, then the change in the energy use of non-participants across the two time periods can provide an estimate of natural conservation. The estimated energy savings for program participation is then reduced by this amount under the assumption that program participants are similar to non-participants and would have reduced their consumption by this same amount, had the program not been in place.

While providing one approach for addressing natural conservation, this procedure is not able to assess the problem of self-selection bias. Self-selection bias can occur when participants are compared with non-participants to obtain estimates of program-induced energy conservation. For example, program participants may be more aware of conservation opportunities and are more conservation conscious relative to non-participants. If this is the case, then using non-participants as the control group would underestimate the amount of natural conservation that would have occurred among the participants had the program not existed. There are several methods which can be used for correcting for self-selection bias.[4] This project uses a method developed by Dubin and McFadden (1984), and the approach was chosen for two reasons: (1) it does not require estimation of any additional equations; and (2) it introduces a new term into the equation which can be used to test whether or not there is indeed self-selection bias.

The statistical analysis undertaken in this project involved utilizing discrete choice techniques to estimate program participation and multivariate regression to estimate program energy savings. The use of a discrete choice participation model in conjunction with energy savings models offers three advantages when examining the impacts on energy consumption by a demand-side conservation program:

- First, this approach can help to mitigate biases introduced by customers self-selecting themselves into the program.
- Second, discrete choice models can provide useful information on the factors that influence customers' decision to participate in a program.
- Third, discrete choice models can provide a procedure by which future program participation can be estimated.

Discrete choice analysis encompasses a variety of different estimation techniques including Logit and probit estimation, and linear probability techniques. It was decided to use a Logit model for this analysis due to the fact that these equations are fairly easy to estimate and the formula for the logit probability is readily interpretable.[5]

The formulation of the discrete choice model begins by noting that the value of participating or not participating can be expressed as:

$$V_j = c_j + Z\varphi_j + w_j, \qquad j = p,n$$

where Z is a vector of such things as expected monetary benefits from participating or not, socioeconomic factors, tastes about energy efficiency, and awareness of the program, V_p is the value of participation in an energy conservation program, and V_n is the value of not participating. An individual will choose to participate if and only if V_p V_n. Since V is not observed (only the choice of participating is observed), probabalistic methods must be used. The probability of participation is:

$$P_p = \text{Prob}\,(V_p + w_p < V_n + w_n)$$

If w_j is distributed logistic, then the probability that the consumer chooses to participate in an energy conservation program is:

$$P_p = \frac{e^{Z\varphi}}{(1 + e^{Z\varphi})}$$

The parameters φ, and thus the probability of participation, can now be estimated by maximum likelihood methods.

The general specification of the participation model differs depending upon which program is being evaluated:

- *HESP Audit and Low/No Interest*: The participation model is a two equation model. The first equation uses discrete choice methods to estimate an equation examining the audit participation decision. The second equation examines the decision to undertake specific actions recommended by the auditor.
- *Seal-up*: The participation model for this program is similar to that for the audit program. The first stage estimates the probability of a given household participating in the program. The second stage estimates the probability that the households that do participate engage in more extensive water heater related conservation measures and/or additional weatherization actions.
- *Appliance Rebate*: Since this program requires a sizable investment on the part of consumers, the factors that affect participation in the rebate program are likely to be very different from those which affect audit or seal-up participation. The model for this program is a single equation logit.
- *Low Income Direct Grant Programs*: These are very different in character from the other three conservation programs. One major difference is that this program is not administered by the utilities. Rather, it is implemented by Community Action Program (CAP) Agencies. This leads to extensive data problems and the nature of the program prevents the estimation of a robust participation model.

Once the participation model is estimated, multivariate regression is used to estimate the energy savings. Unlike simple comparison techniques, multivariate regression approaches are more flexible in their ability to control for factors that influence energy consumption and savings. There are several different specifications of the regression models that can be used to estimate program savings. These can be broadly classified into three categories—standard models, change or difference models, and conditional demand models. The general structure of these models is outlined in Violette and Yokell (1987b). The conditional demand model has been widely used in recent demand forecasting studies and is an excellent technique for this purpose. However, since the purpose of the project is to investigate the change in energy use rather than forecasting, the model of choice for this analysis is a combined conditional demand/change model for reasons described below.

The conditional demand model is derived from the simple identity that the change in total energy consumption must be equal to the change in the consumption of all the energy using equipment in the household. This can be expressed through the following equation:

$$ENERGY_{it} = e_{it0} + e_{it1}A_{i1t} + e_{it2}A_{2t} + \ldots + e_{int}A_{nt} + \varepsilon_{it}$$

where:

$ENERGY_{it}$ = the total energy consumed by household i at time t

A_n = a variable that takes on the value of 1 if the household has that appliance at time t

e_{int} = the energy consumed by appliance N in household i at time t, and

ε_{it} = the error term.

The e_{int} is termed the conditional kWh use of appliance N and it is not necessarily constant over time or over households. Instead, for each customer this term depends upon the characteristics of that customer and external factors such as energy prices. Assuming a simple linear form, this relationship can be expressed as:

$$e_{int} = C_{int0} + \sum (C_{intk} Z_{intk})$$

where Z_{intk} is the k^{th} characteristic of customer i at time t that influences energy use of appliance N, and C_{intk} is the estimated coefficient. These characteristics could include such things as household income, floor space of the house, and whether the household had participated in a utility conservation program. Since it is unusual for there to be metered data on individual end uses across an adequate sample of households, the C_{intk} can be estimated indirectly by substituting the equation for e_{int} into the $ENERGY_{it}$ equation, yielding:

$$ENERGY_{it} = [C_{i0t0} + \sum (C_{i0tk} Z_{i0tk})] + \ldots + [C_{intk} + (C_{int0} Z_{intk})] \times A_{int}$$

where:

$$e_{it0} = [C_{i0t0} + \sum (C_{i0tk} Z_{i0tk})]$$

.

.

$$e_{int}A_{int} = [C_{intk} + \sum (C_{int0} Z_{intk})] \times A_{int}$$

Since the purpose of this project is to estimate the change in energy use resulting from conservation actions, a more useful specification of the above model would be to incorporate the change in energy consumption by including the pre-program energy use as an additional explanatory variable. Thus, the characteristics which have not changed between the two time periods is implicitly incorporated in the pre-program energy use coefficient. If this is incorporated into the conditional demand equation, the model only requires inclusions of variables that contribute to either the change in energy use of appliance N between the two periods (which would include information on changes in energy prices, changes in the number of household members, and program participation), and the change in the appliance holdings of the family between the two periods. Formally, the equation is:

$$\begin{aligned} ENERGY_{it} = {} & [C_{i0t0} + \sum (C_{i0tk} DZ_{i0k})] \\ & + \ldots + [C_{intk} + \sum (C_{int0} DZ_{ink})] \times A_{itn} \\ & + \ldots + [C_{intk} + \sum (C_{int0} Z_{int}k)] \times DA_{in} \\ & + C_e ENERGY_{it-1} + PP \end{aligned}$$

where,

$ENERGY_{it-1}$ is the energy consumed by the household during the previous period, Ce the estimated coefficient, DZ_{ink} is the change in the k^{th} characteristic of customer i between the two periods, DA_{in} is the change in the household's appliance holding for each appliance N between the two periods, and PP is a dummy variable denoting participation in a conservation program.

This formulation provides two significant advantages. First, the estimation is simplified since the only important variables are those that have changed between the time periods. Second, by incorporating the previous period's energy consumption into the equation, the equation will have high explanatory power in terms of overall R^2.

It is important to realize that estimating the basic conditional demand equation for each year and then subtracting them to get the change in energy consumption is not the same as including the previous period's energy demand as an additional explanatory variable. The reason is that by simply subtracting, several terms drop out which probably should not, notably the shift parameters. In addition, subtraction implicitly constrains the parameter on the previous year's energy consumption to equal to one, while the change specification is more general.

The energy demand also includes a "selectivity correction" term which corrects any possible self-selection bias. This term was developed by Dubin and McFadden (1984) and is simply a formula for the mean of ε_{it} given certain distributional assumptions. Since all that is being done is adding a term which equals the expected value of the error term of the (uncorrected) energy demand equation, this new equation is insured to have an error term with a zero expected value, hence it is unbiased.

Engineering Simulation and Additional Analysis

This analysis is currently in the statistical estimation phase. The next steps in the analysis will be an engineering simulation, valuation of energy savings, and then the project will conclude with a cost/benefit analysis. This section outlines these steps.

There are two types of possible energy savings associated with any conservation program: reduction in the total amount of energy consumed, and reductions in peak energy demand. However, billing data is only available for kilowatt-hours or monthly therms. Therefore, there is no statistical method available to estimate peak demand savings. This analysis uses an engineering simulation (ESPRE) to estimate the change in peak and off-peak energy savings. ESPRE provides hourly estimates of residential thermal load and energy use by using building energy analysis. The basic procedure is to use the kilowatt-hour estimation from the regression model in ESPRE to produce load shape estimates. From this, it is then possible to get estimates of the peak reduction.

In addition, the ESPRE will also be used to compare the estimated energy savings from the statistical model and the estimated energy savings from the auditor's report. Comparison of these three estimates and some analysis of the reasons why they differ may prove to be a fruitful exercise.

Another step in this analysis will involve determining a value for the estimated energy savings attributed to the conservation programs. For electricity, in the short run, saved kilowatt-hours can be valued at the short-run marginal cost of producing them. In the long-run, saved kilowatt hours have the potential to reduce the need for new generating facilities. Though an individual conservation program is unlikely to have such a large effect, multiple conservation programs across several utilities such as the ones in this study could result in a significant deferral of the need for new generating facilities.

For valuing gas saved by conservation, it is necessary to determine a dollar value for that gas at different times of the year for many years into the future. The process of determining the value of a unit of gas saved requires an understanding of the gas supply sources and distribution system of each utility, together with the characteristics of the demand faced by each utility.

The final step in this analysis will be a comparison of the estimated benefits and cost resulting from the conservation programs. This will be examined from a number of different perspectives, *i.e.*, utility, customer, participant, and non-participant.

Conclusions

This paper has outlined the procedure used to evaluate a multiple program/multiple organization conservation project in New Jersey. There are several significant and unique aspects to this project:

- The sheer size of the project is unique: the NJCAT project involves seven utilities, and over 50 conservation programs.
- The multi-program/multi-utility nature of the project allows for the pooling of data to improve the precision of energy savings estimates and the comparison of the effectiveness of various programs.
- Data were obtained over a three year period to investigate whether the initial energy savings associated with a program declines over time after participation increases.
- The design of the survey questionnaire was carefully thought out to incorporate questions which gave indications of previously unobservable household characteristics.
- The auditor's report was input into a computer to allow easy access. This enables comparison between the auditor's estimate of energy savings and that found by the analysis.
- Several types of techniques were undertaken to estimate energy savings: a simple comparison using PRISM, a statistical estimation of both a nested discrete choice participation model and a hybrid condition demand/change energy savings model, and an engineering simulation using ESPRE.
- Estimates of both total energy saved and reductions in peak energy demand were found by combining the statistical model with the engineering simulation.

Several estimation issues are addressed in the analysis. These include the problems related to self-selection bias and techniques to control for this, the pros and cons of using PRISM as a pre-processor for the statistical model, and the merits and drawbacks of several different model specifications for determining energy savings.

Endnotes

[1]For a more complete disscussion of the various programs and the evaluation issues raised in this paper, see Violette and Yokell, (1987a).

[2]Assistance eligible customers are low-income and senior citizen customers who participate in federal assistance programs (*e.g.*, Aid to Dependent Children).

[3]Oil and gas burners have the same efficiency if the oil burner tip is clean. Since it is standard practice for houses burning oil to have a maintenance agreement covering routine burner maintenance, the use of gas as a proxy is a reasonable assumption.

[4]See Violette and Ozog (1989) for a discussion of the application of self-selection bias techniques and their implication for program evaluation.

[5]See Train (1986) for a discussion of the methods and relative merits of probit, logit, and other discrete choice estimation techniques.

References

Dubin, J., and D. McFadden. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," ***Econometrica,*** 52. 1984.

Train, K. ***Qualitative Choice Analysis.*** Cambridge, MA: MIT Press. 1986.

Violette, D., and M. Ozog. "Correction for Self-Selection Bias: Theory and Applications," presented at the . . . [?].

Violette, D., and M. Yokell. ***Phase I Final Report: Methods Analysis for the Conservation Program Evaluation Project,*** RCG/Hagler, Bailly, Inc. submitted to the New Jersey Conservation Analysis Team. 1987a.

Violette, D., and M. Yokell. "Estimation of Conservation Program Energy Savings using Customer Billing Data," ***Papers and Proceedings of the Eighth Annual North American Conference, International Association of Energy Economists.*** 1987b.