Measurement of DSM Program Savings: Comparing Estimates from Treatment-Effects and Fixed-Effects Models

Eric J. Solberg, California State University, Fullerton
Andrew M. Gill, California State University, Fullerton
Abdullah Y. Ahmed, Occidental Analytical Group

ABSTRACT

Treatment effects, program saving rates, for two separate program years of a third party Demand Side Management (DSM) Program are estimated by regression using a treatment effects model where variables are measured in levels. These estimates are compared to those obtained from a fixed effects model where variables are measured as changes over time. The fixed effects model is likely to generate estimates that are better suited for the purpose of predicting savings if participants were chosen at random from a population of potential participants rather than opting for treatment because unobserved site characteristics make anticipated program savings higher.

Introduction

This paper compares two least-squares estimation strategies. First, the DSM program treatment effect is measured using a dummy variable coefficient estimated using level data. Second, average treatment savings is measured using a fixed effects model using change (first differenced) data. In both cases panel data are used. One advantage of the fixed effects model is that there is no need to measure nor to control for differences in traits that vary across sites but do not vary over time. Costly on-site surveys therefore, may be unnecessary. Another advantage is that bias on the estimated program savings caused by unobserved heterogeneity across sites is circumvented. Therefore, fixed effects estimation may be cheaper and more reliable.

The California Public Utilities Commission (CPUC) directed California investor-owned utilities to develop and implement pilot bidding programs to assess alternative demand-side management delivery mechanisms. An energy service provider, Energx Controls of Cypress, California, had developed a patented controller (smart thermostat) to regulate water temperature based on usage in recirculating hot water systems in multi-family central water heating and hydronic heating systems. Energx responded to CPUC's directive with a proposal to the Southern California Gas Company to install this system and measure its impact on the consumption of natural gas at multi-family dwellings.

Management firms and owners were given incentives to install the system because this industry typically requires a short payback period for new investments. Controllers were installed in two phases. In the first program period of 1995-96, 340 controllers were installed that affected 14,709 apartment units. In the second program year of 1996-97, 492 controllers were installed that affected 24,905 apartment units. A load impact study was conducted for each of the initial program years, and each was followed by a load impact and retention study in the fourth year after installation.

This paper describes two of the measurement and evaluation models that were used to assess the savings for each of the initial program years. Two alternative least squares regression approaches were used to estimate program savings. First, trimmed least-squares was used to estimate a treatment effects model with and without control group data. Explanatory variables included a dummy variable to indicate whether a site was a participant or control site, a dummy variable to indicate the post-installation period for the treatment group, site characteristics that did not vary over time, variables like weather that did vary over time and across geographical location, and trend. A problem with the treatment effects model is that least-squares parameter estimates may be biased and inconsistent due to unobserved heterogeneity across sites. One source of this unobserved heterogeneity is self-selection when treatment sites choose to participate because of greater anticipated savings. The second model, a fixed-effects or change model, was estimated including control group data. The change model is a cost-effective way of estimating program savings because it does not require information on site characteristics, and the fixed-effects model avoids some of the pitfalls associated with self-selection bias.

Typically, the fixed effects model is estimated as a first-differenced equation. While it is a single cross-sectional equation, each variable is differenced over time. A key assumption is that explanatory variables Δx_i are uncorrelated with the error Δe_i. If this condition is satisfied, the fixed effects coefficient estimators are unbiased. An assumption required for hypothesis testing is that of homoskedasticity, but correction methods are available if heteroskedasticity is present just as in ordinary regression estimation. In some applications, loss of degrees of freedom can be an issue because observations are lost when differencing if the panel is unbalanced or when there are missing observations. In fact, in the application at hand, the number of observations is cut at least in half because differences are calculated over a 12-month displacement length. Therefore, one must begin with a large enough sample, but otherwise the problems are no different than those encountered in any least squares regression exercise.

Methodology

Survey Data

On-site surveys were conducted at participant and control sites to collect information on site characteristics that were likely to influence natural gas consumption. Billing records were used to gather monthly data on therms of natural gas consumed over time for each site 12 months before and up to 12 months after treatment and for control group sites for comparable periods. Monthly rainfall and temperature data were matched with the geographical location of the sites.

Non-treatment comparison sites could not be selected on an entirely random basis. Since the treatment sites were self selecting due to their larger size and existence of central water heaters with re-circulating loops, the control group had to be selected accordingly. Many owners and managers of non-treatment sites were not eager to cooperate – partly because they had to agree to reveal billing data, and partly because they did not want to be interviewed or have inspectors survey their facilities. Nevertheless, 150 sites were identified and used as a control group for the first program year. Most of these sites opted to become treatment sites in the second program year, and so a new set of 152 control sites was identified for the second program year. By the time the persistence studies began, most of the remaining control group sites had chosen to install the system, and there were insufficient numbers of appropriate sites to include control groups in the analyses. Therefore, this paper will focus on the results from the two initial program years.

Regression Models

A specification search was performed as part of the first year impact study for the first program year, and the specification was refined during the first-year impact study of the second program year. As an initial step, a fully interactive quadratic form was estimated as a second-order approximation to any underlying non-linear functional form. As might be expected, severe multicollinearity was detected, and a more parsimonious (and more logical) specification was selected. While some collinearity undoubtedly remains, coefficient estimates are not biased by multicollinearity. Definitions of variables used in the analyses are reported in Table 1. The variables are measured as monthly levels. THERMS is the number of therms per month used per account. There is some measurement error here since the number of billing days varied somewhat each month. Since there could be more than one meter at each site, a weight in the unit interval was created called FACTOR that equals the therms per meter divided by the total therms per account. As an example, if there are 3 meters each measuring 1/3 of the total usage on a billing account, FACTOR is equal to 1/3. ADJTHRMS measures the average monthly level of consumption for apartment units served by each meter per apartment unit, ADJTHRMS = FACTOR*THERMS/UNITS. The other variables listed in Table 1 are self explanatory with the exception of PROBLEM. This dummy variable indicates whether a problem like a slab leak occurred that may have affected the efficiency of the system. Table 2 reports the means of the variables for the treatment sites and control sites for both program years.

Treatment Effect (Level) Model The treatment effect, or dummy variable, model of conditional demand estimated is:

$$\ln y_{it} = \mathrm{a} + \mathbf{b}\boldsymbol{z}_i + \mathbf{c}\boldsymbol{x}_{it} + \mathrm{d}T_{it} + e_{it} \tag{1}$$

The variable $\ln y_{it}$ represents the natural logarithm of ADJTHRMS consumed for the i^{th} controller-site during month t. The vector $\boldsymbol{z}_i$ represents a collection of site-specific characteristics that vary by site but do not vary over time. The vector $\boldsymbol{x}_{it}$ represents a collection of variables, such as measures of weather, which vary over time as well as by site. The term e_{it} represents a random error. The variable T represents the treatment (installation and operation of the controller) that is both time and site-dependent. When this model is estimated using only installation-site data, the treatment variable T is time-dependent only. T is equal to zero prior to installation, and T is equal to unity after a post-installation adjustment period.

One advantage of measuring the dependent variable in logarithms is that the estimated coefficient for d, which is expected to be negative, can be used to calculate an estimate of the percentage savings using the following formula:

$$\text{savings percent} = 100\{\exp(-\mathrm{d}) - 1\}$$

An additional advantage is that the distribution of log therms is more symmetric than the distribution for therms. The distribution of therms is skewed right because of a lower limit of zero and no bound on the upper limit – empirically, the right-hand tail is longer than the left-hand tail. Moreover, the fit of the estimated equation is much better when the logarithm is used, and the estimated treatment effect is more precise but not significantly different in magnitude.

Table 1 Variable Definitions

Variable	Definition
THERMS	Therms per month per account
FACTOR	Therms per meter divided by total therms per account
UNITS	Number of apartment units served by the boiler
ADJTHRMS	= Factor∗Therms/Units
LNADJTHM	= ln(ADJTHRMS)
ISSITE	= 1 if installation (treatment) site; else = 0
POSTIS	= 1 if treatment site after installation; else = 0
PROBLEM	= 1 if a problem occurred that affected consumption; else = 0
HOA	= 1 if home owners association; else = 0
STORS	number of stories
SOLAR	= 1 if solar water heating is present; else = 0 (number per unit in 2nd year)
DRYRS	Number of centrally located gas clothes dryers (per unit in 2nd year)
BBQS	Number of gas barbeques (per unit in 2nd year)
HTDPOOL	= 1 if a heated pool is present; else = 0
SPA	= 1 if spa is present; else = 0
POOLSPA	= 1 if pool and spa is present; else = 0
HYDRNIC	= 1 if hydronic space heating; else = 0
CONVERT	= 1 if converted from hydronic heating; else = 0
HTRBTU	Heater BTU (per unit in 2nd year)
SPHEAT	= 1 if space heating is available; else = 0
GASCKG	= 1 if apartments have gas cooking; else = 0
RAINFALL	Rainfall in inches by month by region
MINTEMP	Maximum temperature by month by region
MAXTEMP	Minimum temperature by month by region
OCCRATE	occupancy rate as a percent
MONTH	A time trend starting at Jan-1994 = 1

Table 2 Variable Means

	First Program Year		Second Program Year	
Variable	**Installation Sites**	**Control Sites**	**Installation Sites**	**Control Sites**
ADJTHRMS	36.526	33.080	29.740	39.480
PROBLEM	0.014	0.000	0.098	0.000
HOA	0.205	0.080	0.100	0.133
STORS	2.102	1.953	2.299	1.943
SOLAR*	0.126	0.000	0.001	0.004
DRYRS*	2.847	3.007	0.031	0.014
BBQS*	0.163	0.127	0.005	0.004
HTDPOOL	0.125	0.147	0.157	0.129
SPA	0.143	0.113	0.137	0.089
POOLSPA	0.073	0.107	0.118	0.085
HYDRNIC	0.329	0.193	0.193	0.298
CONVERT	0.014	0.000	0.000	0.000
SPHEAT	0.064	0.153	0.232	0.205
GASCKG	0.072	0.107	0.231	0.186
RAINFALL	1.521	1.444	1.645	1.629
MINTEMP	54.876	55.379	54.987	55.167
MAXTEMP	75.905	73.417	75.463	72.802
OCCRATE	93.663	94.744	NA	NA

* Variable defined differently in second program year.

The treatment effects model can be estimated by including the control group for comparison with an additional indicator for whether a site is a treatment site.

$$\ln y_{it} = \mathrm{a} + \mathbf{b}\boldsymbol{z}_i + \mathbf{c}\boldsymbol{x}_{it} + \mathrm{d}_1 T_{1it} + \mathrm{d}_2 T_{2it} + e_{it} \qquad (2)$$

In this specification, T_1 indicates that a site is a treatment site (SITE), and T_2 indicates the post-treatment period (POSTIS). The estimate of d_2 is used to measure the treatment effect.

Fixed Effects (Change) Model It is likely that sites differ in characteristics that are not observed or may not be observable, but which nonetheless affect consumption. The problem is one of unobserved heterogeneity. This may lead to statistical estimates of savings based on ordinary least squares that could be expected at non-participant sites that are biased and inconsistent. Such bias occurs when the unobserved characteristics that affect consumption are correlated with explanatory variables included in the regression equation creating an omitted variables bias. One manifestation of this problem is the so-called model of self selection. In the present problem, self selection could arise if treatment sites opted for treatment because, for reasons unobserved by the analyst, they anticipated greater savings.

A solution to this problem is to estimate savings using a fixed-effects model (Greene, 2003; Wooldridge, 2003). The fixed effects model exploits the panel aspect of the data to analyze how changes in the level of the dependent variable are related to changes in the levels of the independent variables. If the unobservable characteristics are fixed over the period of analysis, the effects of any unobserved heterogeneity are eliminated.

Consider the following specification

$$y_{it} = \mathrm{a} + \mathbf{b}\boldsymbol{z}_i + \mathbf{c}\boldsymbol{x}_{it} + \mathrm{d}T_{it} + f_i + e_{it} \tag{3}$$

Observable fixed effects are measured by the $\boldsymbol{z}_i$ vector, and unobservable fixed effects are captured by f_i. At some later period $t+m$ where m represents some displacement period, the equation becomes

$$y_{it+m} = \mathrm{a} + \mathbf{b}\boldsymbol{z}_i + \mathbf{c}\boldsymbol{x}_{it+m} + \mathrm{d}T_{it+m} + f_i + e_{it+m} \tag{4}$$

If the displacement length m is chosen so that $T_{it} = 0$ before treatment and $T_{it+m} = 1$ after treatment, subtracting (3) from (4) results in

$$y_{it+m} - y_{it} = \mathbf{c}(\boldsymbol{x}_{it+m} - \boldsymbol{x}_{it}) + \mathrm{d}(T_{it+m} - T_{it}) + (e_{it+m} - e_{it}) \tag{5}$$

Now the change in consumption over time is related only to changes in variables that change over time. The effects of site characteristics that do not vary over time cancel out of the change equation. The variable $(T_{it+m} - T_{it})$ equals unity after the treatment period, and otherwise the dummy variable equals zero. Since the displacement length is chosen so that change in the variable is measured between an after-treatment observation and a before-treatment observation, the estimated coefficient for d provides a measure of the reduction in adjusted therms due to the treatment. If one wants to allow for the possibility that some of the fixed effects vary over time, one can estimate the model with an intercept.

Results

Treatment Effects (Level) Model

The parameter estimates for the treatment effects model are reported in Table 3. White's technique (1980) was used to correct for any heteroskedasticity. Probability values are reported for a two-tailed test of statistical significance of the coefficients. Trimmed least squares was used to eliminate observations where the residuals in a first-step regression fell outside a 95% percent confidence interval. The application of trimmed least squares did not change the estimated treatment effect in any appreciable way except to make it more precise (a smaller standard error). There was no evidence of any serial correlation once weather variables were included as explanatory variables. Observations on the consumption of therms were treated as missing values during the month of installation and, in a few cases, during a subsequent period when the system was fine tuned.

Table 3 Trimmed least squares estimates of log(ADJTHRM), treatment effects model.

	1st program year, no control group		1st program year, with control group		2nd program year, no control group		2nd program year, with control group	
VARIABLE	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
INTERCEPT	5.9524	0.0000	6.2789	0.0000	4.4673	0.0000	4.9688	0.0000
SITE			0.1943	0.0000			-0.2834	0.0000
POSTIS	**-0.1630**	0.0000	**-0.1581**	0.0000	**-0.2401**	0.0000	**-0.1968**	0.0000
PROBLEM	0.2538		0.2209	0.0000				
HOA	0.0219	0.0082	0.0159	0.0363	-0.1954	0.0000	-0.2200	0.0000
STORS	-0.3010	0.0000	-0.3721	0.0000	-0.1271	0.0000	-0.1395	0.0000
SOLAR*	-0.0053	0.0000	-0.0049	0.0000	-1.3552	0.0000	-1.6113	0.0000
DRYSU*	-0.0196	0.0000	-0.0169	0.0000	-0.4416	0.0000	-0.4358	0.0000
BBQNUM*	0.0230	0.0000	0.0003	0.9380	-0.9705	0.0000	-0.5283	0.0001
HTDPOOL	0.2149	0.0000	0.2392	0.0000	0.1042	0.0000	0.0341	0.0097
SPA	0.0769	0.0000	0.0368	0.0046	-0.0465	0.0478	-0.0807	0.0030
POOLSPA	-0.1418	0.0000	-0.1405	0.0000	-0.1541	0.0000	-0.0534	0.0489
HYDRNIC	0.1875	0.0000	0.1620	0.0000	0.1519	0.0000	0.0789	0.0000
CONVERT	-0.1924	0.0000	-0.2766	0.0000				
HTRBTU*	0.0012	0.0000	0.0014	0.0000	0.0000	0.0000	0.0000	0.0000
SPHEAT	0.1819	0.0000	0.3482	0.0000	0.3256	0.0000	0.2804	0.0000
GASCKG	-0.0242	0.1122	-0.0250	0.0524	0.1539	0.0000	0.1864	0.0000
RAINFALL	0.0476	0.0000	0.0372	0.0000	0.0635	0.0000	0.0587	0.0000
RAINFALL^2	-0.0031	0.0000	-0.0023	0.0000	-0.0036	0.0000	-0.0035	0.0000
MINTEMP	-0.0061	0.0000	-0.0101	0.0000	-0.0116	0.0000	-0.0078	0.0000
MAXTEMP	-0.0204	0.0000	-0.0181	0.0000	-0.0114	0.0000	-0.0158	0.0000
OCCRATE	-0.0018	0.1309	-0.0053	0.0000				
MONTH					0.0034	0.0000	0.0017	0.0000
R-squared	0.4757		0.4539		0.4879		0.5368	
Mean Dep Var	3.3757		3.3716		3.1924		3.2957	
Sample size	11,439		16,519		16,957		22,202	

* These variables were defined differently between the first and second program years.

Refinements were made to the model specification between the first and second program years, and so some of the explanatory variables were measured differently between the two program years. In the first program year, the dummy variable PROBLEM indicates that a serious problem occurred that may have affected the consumption of therms. One case, where a maintenance man repeatedly tampered with the system's settings, was referred to as "the case of the maintenance man from hell." The exclusion of observations from this site made no perceptible difference in the other parameter estimates. The variable SOLAR is a dummy variable in the first program year indicating the presence of solar water heating, but SOLAR equals the number of solar heating installations per apartment unit at the site for the second program year. The variables DRYRS and BBQS measure the numbers of centrally located gas dryers and barbeques at each site for the first program year, but DRYRS and BBQS are divided by the number of units at the site for the second program year. In the first program year, HTRBTU is the BTU rating of the water boiler at the site, but in the second program year that rating was divided by the number of apartment units served by the boiler. In the first program year, the variable OCCRATE was used to control for changes in occupancy rates across sites and over time during the survey period, but that variable had insufficient variation during the second survey period – the occupancy rates were high and not changing. Therefore, to control for other changes in the economic environment, a monthly time trend MONTH was added to the model as a proxy.

The coefficients of primary interest are for the installation site treatment indicator POSTIS. Those parameter estimates are indicated in bold face in Table 3. Clearly there are substantial program savings due to treatment.

Fixed Effects (Change) Model

Changes in therms are calculated as follows. For each treatment site, subtract pre-treatment therm consumption for a particular month from post-treatment consumption for the same month one year later. If there is a reduction in consumption, the resulting number will be negative. Analogous calculations are made for the comparison group sites. Since there is no treatment effect for the non-installation sites, the expected change over the period is zero if other influences are constant.

Equation (5) is estimated by ordinary least squares. The only observable variables that change are the weather variables, MINTEMP, MAXTEMP and RAINFALL. An intercept term is included in equation (5) to allow for the possibility that some unobservable effects change over time. Savings are estimated using two strategies. In the first, the parameters are estimated for a pooled sample that includes sites in both the treatment and control groups. In that model, the coefficient of interest is that for ISSITE. Since it is expected that reductions in consumption will be larger or increases in consumption will be smaller at the treatment sites, the expected sign for the coefficient of the treatment effect dummy is negative. Moreover, the magnitude of the coefficient will give a direct estimate of average savings. For example, suppose that the coefficient estimate is –5. This suggests that, on average, the change in adjusted therm consumption at treatment sites is 5 therms lower than at comparison sites, controlling for changes in temperature and rainfall.

One limitation of the above approach is that it restricts the coefficients of the time and region-varying measures to be the same across treatment and comparison groups. An alternative is to relax this restriction and estimate the parameters of Equation (5) separately for the two groups. Again, an intercept is added to each equation to account for possible unobservable effects that vary over time. Once the parameter estimates are obtained, savings are estimated by applying the parameter estimates for the treatment sites to the average characteristics of the comparison sites. This generates an estimate of what the average change in the level of savings would have been if the treatment had been applied to the control sites. As a final step, this predicted change in therm consumption is compared to the actual change in therm consumption for the control sites. The difference between these magnitudes is an estimate of savings.

The results for the fixed effects (change) models for the first program year are reported in Table 4:

Table 4 Fixed effects models for 1st program year

	Pooled sample		Installation sites		Control group	
Variable	Coef.	p-value	Coef.	p-value	Coef.	p-value
Intercept	0.2647	0.0131	**–4.8642**	0.0001	**0.1825**	0.0695
ISsite	**–5.1313**	0.0001				
ΔRainfall	–0.2454	0.0001	0.2133	0.0004	0.2650	0.0001
ΔMintemp	–0.2906	0.0001	–0.2506	0.0001	–0.3404	0.0001
ΔMaxtemp	–0.1837	0.0001	–0.3154	0.0001	–0.0607	0.0001
R-squared	0.1688		0.0624		0.0415	
Mean Dep Var	–2.3163		–5.0318		–0.1205	
Sample size	6,511		2,911		3,600	

The coefficient for the variable ISSITE, indicated in bold type in the table, is an estimate of savings from the treatment. The coefficient estimate of –5.131 is interpreted as follows. Controlling for changes in weather, the change in consumption at treatment sites is –5.131 adjusted therms lower on average than at control sites. For the second program year, the average therm consumption at treatment sites is 33.08 adjusted therms prior to treatment. Thus, the treatment effect indicates a 15.5 percent reduction in consumption for the second program year.

Turning now to the results of the model when separate equations are estimated for the treatment and control groups, the estimated intercepts are important. Each intercept is included to capture any time-varying characteristics not included in the model. For the treatment group, this measures the effect of the treatment. The estimated intercept indicates that treatment reduced average consumption of adjusted therms by –4.86 units in the post-treatment period. For the comparison group, average adjusted therm consumption rose by 0.1825 units over the same period. As a final step, the change in average consumption is predicted for the comparison group using the parameter estimates for the treatment group. This procedure yields an estimate of what the change would have been for the comparison group had they had the treatment. Using the means of the changes in the weather variables, the average change in consumption is given as:

$$\text{Predicted change} = -4.86 + (-0.315*(0.898) - 0.251(0.407) + (-0.213)) = -5.33$$

The mean change in consumption for the comparison group is –0.12. Thus, the average change in adjusted therm consumption would have been 5.21 therms lower had the comparison group received the treatment. This yields a percentage reduction of 15.7 percent, which is slightly higher than the estimate obtained from the ISSITE coefficient from the pooled data.

The results for the fixed effects (change) models for the second program year are reported in Table 5:

Table 5 Fixed effects model for 2nd program year

	Pooled sample		Installation sites		Control group	
Variable	Coef.	p-value	Coef.	p-value	Coef.	p-value
Intercept	0.8821	0.0001	**–3.0971**	0.0001	**0.8917**	0.0001
ISsite	**–4.0151**	0.0001				
ΔRainfall	0.3095	0.0001	0.2890	0.0001	0.3198	0.0001
ΔMintemp	–0.4944	0.0001	0.4858	0.0001	–0.4905	0.0001
ΔMaxtemp	–0.0891	0.0001	–0.0426	0.0762	–0.1658	0.0001
R-squared	0.1826		0.0772		0.1399	
Mean Dep Var	–0.7609		–2.8380		1.1105	
Sample size	7,507		3,558		3,949	

The coefficient for the variable ISSITE, indicated in bold type in the table, is an estimate of savings from the treatment. The coefficient estimate of –4.015 is interpreted as follows. Controlling for changes in weather, the change in consumption at treatment sites is –4.015 adjusted therms lower on average than at control sites. For the second program year, the average therm consumption at treatment sites is 29.74 adjusted therms prior to treatment. Thus, the treatment effect indicates a 13.5 percent reduction in consumption for the second program year.

Turning now to the results of the model when separate equations are estimated for the treatment and control groups, the estimated intercepts are important. Each intercept is included to capture any time-varying characteristics not included in the model. For the treatment group, this measures the effect of the treatment. The estimated intercept indicates that treatment reduced average consumption of adjusted therms by –3.10 units in the post-treatment period. For the comparison group, average adjusted therm consumption rose by 0.8917 units over the same period. As a final step, the change in average consumption is predicted for the comparison group using the parameter estimates for the treatment group. This procedure yields an estimate of what the change would have been for the comparison group had they had the treatment. Using the means of the changes in the weather variables, the average change in consumption is given as:

$$\text{Predicted change} = -3.1 + (-0.043*(0.151) - 0.4868(-0.299) + (-0.304)) = -3.05$$

The mean change in consumption for the comparison group is 1.11. Thus, the average change in adjusted therm consumption would have been 4.16 therms lower had the comparison group received the treatment. This yields a percentage reduction of 14.0 percent, which is slightly higher than the estimate obtained from the ISSITE coefficient from the pooled data and slightly lower than savings reported for the first program year.

Table 6 presents a summary of the overall results. Estimates from the fixed effects model for the savings rates are lower than those from the treatment effects model for both program years. The estimates of the actual program savings from the treatment effects model are probably fairly good estimates of actual program savings even though they may be biased upwards slightly. For the first program year, there were 14,709 apartment units affected by the program, and the estimated reduction in therms on an annualized basis is 997,782 therms. For the second program year there were 24,905 apartment units affected, and the estimated reduction in therms on an annualized basis is 1,932,685 therms. Nevertheless, if the program were extended to randomly selected sites, the fixed effects model estimates are likely to be a better indicators of possible savings. For the first program year that model predicts a reduction of 902,838 therms would have been realized on an annualized basis, and for the second program year there would have been a reduction in 1,242,760 therms. Using either set of estimates, the program was cost effective. In fact, the contractor is willing to install the controller without cost if the owner of an apartment site will agree to pay the contractor the cost savings in reduced therms for the first year.

Table 6 Summary of Results

	1st Program Year*	2nd Program Year**
Treatment effects model		
Savings rate	17.13%	21.75%
Average therms saved per year	67.83	77.60
Annual savings	997,782	1,932,685
Fixed effects model		
Savings rate	15.50%	14.00%
Average therms saved per year	61.38	49.90
Annual savings	902,838	1,242,760

* Calculation based on 14,709 units and 33.0 therms consumed per unit per month.
** Calculation based on 24,905 units and 29.7 therms consumed per unit per month.

Conclusion

For the first program year, the treatment effects model resulted in an estimated savings rate of 17.70% when data for participants only were used and 17.13% when data from a control group were included. For the second program year, the treatments effects model resulted in an estimated savings rate of 27.14% when data for participants only were used and 21.75% when data from a control group were included. These are likely to be good estimates of actual program saving rates for the treatment sites for each program year, but they are not good estimates for anticipated savings if the program were extended to sites chosen at random from a population of potential installation sites.

The fixed-effects model estimates of program savings are lower. The fixed-effects model generates an estimated savings rate of 15.50% for the first program year and 14.00% for the second program year installations. These are likely to be better estimates of potential savings if the program were extended to sites chosen at random from the population and not subject to self selection. The lower estimates from the fixed effects models are consistent with the expectation that sites with higher expected saving rates are more likely to opt for treatment. Moreover, the use of a fixed effects model avoids the high cost of collecting and recording data on site characteristics since all necessary data can be drawn from billing records or published weather data.

References

Greene, W. H. 2003. "Section 13.3 Fixed Effects," *Econometric Analysis*, Fifth Edition. Prentice Hall, New Jersey

White, H. 1980. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica* 48:817-838.

Wooldridge, J. M. 2003. "Chapter 14 Advanced Panel Data Methods," *Introductory Econometrics: A Modern Approach,* Second Edition, South-Western, Ohio.