A Matter of Size: Using Ordinary Least Squares to Estimate Energy Savings in the Commercial Sector

Hossein Haeri, Ph.D. and Matei Perussi, Quantec, LLC, Portland, OR
Iris Sulyma, BC Hydro, Vancouver, B.C.

ABSTRACT

Regression analysis is one of the most common statistical techniques used for measurement and verification of savings from energy-efficiency programs. Using panel data (cross-sectional time series), evaluators have used regression models to derive estimates of gross and/or net savings in a variety of research designs and specifications. This paper is about the application of regression analysis for assessing savings in the commercial sector. Using data from BC Hydro's Power Smart Partners Program, it demonstrates the effects of large variances in annual consumption on parameters derived by Ordinary Least Squares (OLS) and the accompanying problem of *heteroscedasticity*. Using visual inspection of data and statistical tests, the paper examines the presence and extent of heteroscedasticity and applies two alternative methods to correct for the problem: 1) assigning a unique intercept to each facility (fixed-effects model); and 2) by transforming all to deviations from individual facility means (deviation model). The results show both models' specifications are effective means of correcting for the problem, though, depending on the heterogeneity of the participant population, some tolerable degree of heteroscedasticity may remain. The one disadvantage of the fixed-effects model is that, by incorporating unique intercepts for each facility, it significantly limits the degrees of freedom.

Program Background

BC Hydro's demand-side management (DSM) programs are planned, implemented and managed through the Power Smart organization. The Power Smart Partner Program (PSP) is the organization's flagship program. It has been one of the largest, most active and successful demand-side management initiatives in North America and is regarded as one of the best energy efficiency programs in its class. Through a unique alliance arrangement with vendors and suppliers and partnering agreements with commercial, governmental and industrial customers, PSP offers a combination of energy-efficiency resource acquisition and market transformation services. It relies on education, technical assistance, financial incentives and credits for self-directed energy management practices to bring about lasting change in the market. PSP participants are also offered an "e-Points" bonus, which is designed to recognize and reward the largest commercial and industrial customers who achieve a minimum of 5% in aggregate electrical efficiency improvement across all of their accounts.

Since its launch in April 2002 through March 2005, PSP had succeeded in establishing over 460 partnerships that have produced nearly 700 GWh of cost-effective savings, exceeding the program's target by over two folds at approximately 80% of the expected cost. The program's annual savings goals through 2011 are expected to increase to a level equivalent to nearly one-half of the identified achievable potentials in BC Hydro's non-residential sector. Net evaluated energy savings for the PS Partner business program, based on the analyses described here, were 210 GWh/yr as of March 2005. By March 2007 the PS Partner commercial program reported 320 GWh/yr net energy savings. This paper describes some of the lessons learned in assessing the impacts of the program using regression techniques to analyze panel data.

Methodology Overview

Regression analysis is a staple of impact evaluation techniques. It has been widely (and effectively) used in analyses of consumption history (billing analysis) to estimate savings from energy-efficiency programs. Evaluators have used ordinary least squares (OLS) regression techniques, typically with panel data, to estimate "gross" and/or "net" savings from conservation programs using a broad range of specifications, including conventional demand analysis, "conditional" demand analysis, and combined statistical and engineering (SAE) models, among others.

The large body of energy-efficiency impact evaluation research suggests regression analysis is generally more effective and OLS is likely to produce better and more reliable results in the residential and small commercial sectors. This is mainly because of greater homogeneity in these populations and that annual consumption tends to be distributed normally, with a relatively small variance. In the commercial sector, on the other hand, where large variations in annual consumption—and annual savings—are expected, OLS estimators tend to be less reliable (or efficient) as the variance in the regression residuals tend to increase with annual consumption—a condition known as *heteroscedasticity*.[1]

Constant error variance, or *homoscedasticity*, is one of the basic assumptions underlying conventional normal linear regression models.[2] The assumption simply means the regression error terms (regression residuals) are normally distributed with a constant variance, that is: $Var(\varepsilon_j) = \sigma^2$ for all j. If the error terms do not have constant variance, they are said to be *heteroscedastic*. The problem with heteroscedastic disturbances is OLS estimation places a greater "weight" on observations which have larger error variances than on observations with small ones. Because of this implicit weighing, OLS parameter estimates remain unbiased and consistent, but they may no longer be "efficient."[3]

Errors may increase as the value of the dependent variable increases. The effect of annual income on discretionary expenditures (such as vacations) is the classical example of heteroscedasticity. Families with high incomes tend to spend more on vacation simply because they have greater discretionary incomes; more importantly, there will be greater variability among such families. In the case of large commercial buildings, error terms associated with very large facilities might have larger variances than error terms associated with smaller ones, and larger facilities might have more variability in their energy consumption—and potential savings. In addition to normal variations in the dependent variable, the problem may also arise as a result of a sampling strategy or measurement error.

Several methods are available to correct for heteroscedasticity. Clearly, better sample stratification is the simplest way to remove (or reduce) heterogeneity in the sample. When using panel (cross-sectional/time-series) data, alternative regression specification may also be used to address heteroscedasticity. These approaches rely on different estimation procedures that fall in the category of Generalized Least Squares (GLS). In some cases, large variations in the independent variable may be controlled by simply transforming the variable (e.g., using energy consumption per square foot instead of total consumption) or by expressing the values of variables in terms of deviations from the mean. It is also possible to apply a weighted least squares estimator, where observations expected to have error terms with large variances are given a smaller weight than observations thought to have error terms with

[1] Note that the problem is simply more acute – and by no means limited to - the large commercial sector.

[2] Discussions of heteroscedasticity, its causes, and methods for correcting it are found in all econometric text books. This description is mainly based on discussion of the topic in Robert S. Pindyck and Daniel Rubinfeld, Econometric Models and Economic Forecasts, Second Edition, McGraw-Hill, 1981, Chapter 6; and Judge, George G. et al, The Theory and Practice of Econometrics, Second Edition, John Wiley & Sons, 1985, Chapter 11.

[3] Recall that in the classical normal linear regression models estimated parameters are assumed to be *unbiased* and *efficient*. Since variances of error terms do not play a role in establishing absence of bias in least square estimates, heteroscedasticity does not affect the bias of the estimated parameters.

small variances. Finally, one might use a "fixed-effects" model, in which each observation is assigned a unique intercept, thus capturing some of the variation in total consumption.

In the present analysis, two alternative approaches were tested through an analysis of covariance to reduce "within" facility fluctuations in consumption by: 1) assigning a unique intercept to each facility (fixed-effects model); and 2) transforming all to deviations from individual facility means (deviation model).[4]

The Data

To explore the appropriateness and effectiveness of various procedures to correct for heteroscedasticity, panel data were assembled for 139 completed large commercial customers. One of the primary objectives in the evaluation was to examine the effectiveness of particular groups of measures through estimating unique savings realization rates for each. An SAE model was thus the appropriate specification**.**

Three alternative specifications were used. In all three cases, average daily electricity consumption during each billing cycle was set as the dependent variable, and engineering estimates of expected savings for lighting and other energy-efficiency measures, and heating degree days were the primary covariates. Preliminary screening of the data revealed wide variation in the dependent variable. Total annual consumption across facilities ranged from 8,000 kWh to 15,300,000 kWh. In the 1^{st} quartile of cases, total usage ranged from 8,000 to 200,000 kWh, and, in the 4^{th} quartile, it ranged from 1,300,000 to 15,300,000 kWh. Therefore, it was expected heteroscedasticity would in all likelihood be a problem. The analysis began with a conventional SAE model with the following specification:

1) Ordinary least squares: $ADC_{it} = \alpha + \beta_1 LIGHTINGEE_i + \beta_2 OTHEREE_i + \lambda_1 HDD_{it} + \varepsilon_{it}$

Where, for each facility i and calendar month t,

- ADC_{it} is the average daily kWh consumption during the pre- and post-participation periods for participants
- α is the intercept
- $LIGHTINGEE_{it}$ is the initial engineering estimate of lighting savings, appears in the estimation only during the post-participation period, and takes on the value of 0 in the pre-program period
- $OTHEREE_{it}$ is the initial engineering estimate of other measure savings, appears in the estimation only during the post-participation period, and takes on the value of 0 in the pre-program period
- β_1 and β_2 represent savings realization rates for lighting and other measures, respectively; a value of -1 represents a 100% realization rate
- HDD_{it} is average daily heating degree days based on facility location
- ε_{it} is the error term

The model coefficients and standard statistics for the OLS model are summarized in Table 1. As can be seen for the standard regression model, the model fit is weak with an R^2 of 0.17. The

[4] It is important to note that the fixed-effects and the deviation models are mathematically identical. A proof of this is found in William H. Greene, Econometric Analysis, Macmillan, 1993 pp. 466-468.

LIGHTINGEE coefficient of 3.66 indicates the program realized 366% of the lighting measure engineering estimated savings. Similarly, the coefficient for the OTHEREE variable is estimated at 2.8, indicating the program realized 280% of estimated savings for the "other" measure category. The weather coefficient has the wrong sign. Clearly, in addition to its overall poor performance, estimated parameters do not even carry the correct signs in this model.

Table 1: Ordinary Least Square (Model 1) Estimation Results

Source	Analysis of Variance				
	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	27304847925	9101615975	234.26	<0.0001
Error	3325	129184800000	38852578		
Corrected total	3328	156489700000			
Root MSE	6233.18363		R-Square	0.1745	
Dependent Mean	3800.25688		Adj. R-Square	0.1737	
Source	Parameter Estimates				
	DF	Parameter Estimates	Standard Error	t value	Prob. t
Intercept	1	2903.08334	196.55147	14.77	<0.0001
LIGHTINGEE	1	3.66044	0.14045	26.06	<0.0001
OTHEREE	1	2.782	0.51565	5.4	<0.0001
HDD	1	-1.33581	10.09369	-0.13	0.8947

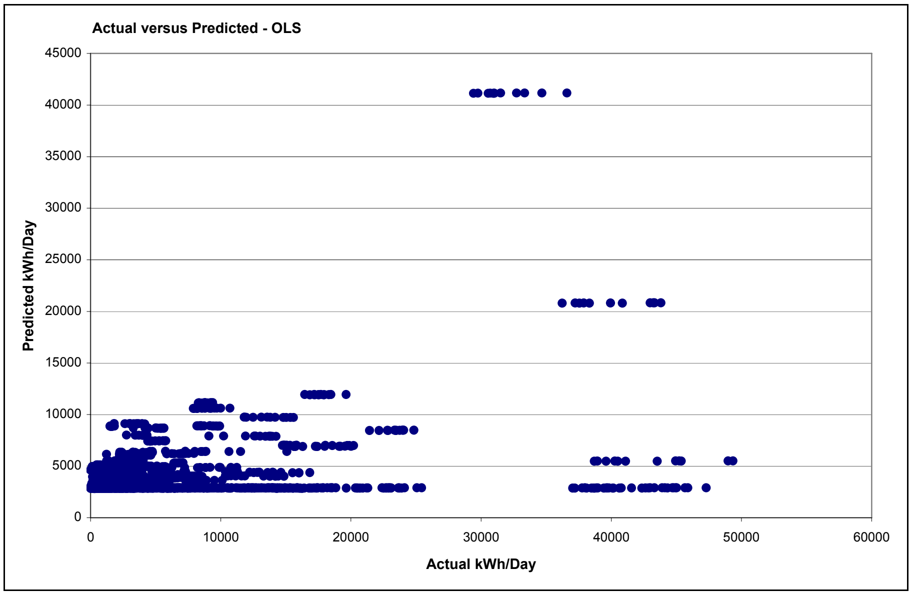
An examination of the plot of actual values against predicted values for this model (Figure 1) clearly demonstrates the inaccuracy of the predicted values. As shown, the predictions also "fan-out" as usage increases. Similarly, the model is plagued by differences in predictions even for the same level of actual consumption. The distribution of residuals in this model clearly indicates a strong relationship between the dependent variable and regression residuals, consistent with the presence of heteroscedasticity. The presence of strong heteroscedasticity was corroborated through the application of the White test. See end notes for a complete description of the White test. Given the overall poor performance of the OLS model, two variations of the GLS of estimators were used to improve the performance of the model and to address heteroscedasticity. The two specifications are shown below.

2) Fixed-effects model: $ADC_{it} = \alpha_i + \beta_1 LIGHTINGEE_{it} + \beta_2 OTHEREE_{it} + \lambda_1 HDD_{it} + \varepsilon_{it}$

3) Deviation model: $ADC^*_{it} = \beta_1 LIGHTINGEE^*_{it} + \beta_2 OTHEREE^*_{it} + \lambda_1 HDD^*_{it} + \varepsilon_{it}$

In the fixed effects model, each facility is treated as a separate case with its own unique intercept (α_i). In model three, the asterisk indicates all variables are converted to deviations from their means; that is, for each variable V in facility i, V_i, $V_i^* = V_i -$ (Mean V_i). Given the differences among facilities with respect to customer sector, facility type and size, and measures installed, this technique better captures the unique characteristics of individual sites that affect electricity use and, hence, savings.

Figure 1. Comparison of Actual and Predicted Values for Daily Consumption from OLS



The results summarized in Table 2 show a significant improvement in the estimated parameters; all now have the correct signs and are statistically significant. (Note that high R^2 of 0.99 does not reflect a strong performance in this case. Indeed, the conventional interpretation of R^2 does not apply because there is no intercept in the model. It is also important to note this model produces 139 intercepts; however, only the average of these is reported.) The coefficients in this model represent savings realization rates for each measure group. For example, the coefficient of -0.77 on the LIGHTINGEE variable indicates 77% of the expected savings from lighting measures were realized.

Table 2: Fixed Effects (Model 2) Estimated Parameters

Source	Analysis of Variance				
	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	142	202204400000	1423974717	1920.92	<0.0001
Error	3187	2362519787	741299		
Corrected total	3329	204566900000			
Root MSE	860.9872		R-Square	0.9885	
Dependent Mean	3800.25688		Adj. R-Square	0.9879	
Source	Parameter Estimates				
	DF	Parameter Estimates	Standard Error	t value	Prob. t
Intercept	139	3844.831936	178.662192	21.03	<0.0001
LIGHTINGEE	1	-0.77117	0.02649	-29.11	<0.0001
OTHEREE	1	-0.90925	0.09623	-9.45	<0.0001
HDD	1	11.39741	1.45949	7.81	<0.0001

Estimation results for the deviation model (3) are reported in Table 3. As in the fixed-effects model, the R^2 is not directly interpretable. Also, note that the mean of the dependent variable now equals 0. It should be noted that all estimated parameters have identical values to the fixed-effects model (2). The plot of actual versus predicted for models 2 and 3, as shown in Figure 2, indicate a significantly better predictive performance compared to the OLS model. As evident in the scatter plot, the "fanning" pattern in the distribution of the residuals is also markedly diminished, and the error terms appear to be distributed evenly around zero (Figure 3). The White test results also suggest an improvement in heteroscedasticity (see endnote 2).

Table 3: Deviation Model (Model 3) Estimated Parameters

Source	Analysis of Variance				
	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	774112733	258037578	363.27	<0.0001
Error	3326	2362519787	710319		
Corrected total	3329	3136632521			
Root MSE	842.80403		R-Square	0.2468	
Dependent Mean	-5.46E-16		Adj. R-Square	0.2461	
Source	**Parameter Estimates**				
	DF	**Parameter Estimates**	**Standard Error**	**t value**	**Prob. t**
LIGHTINGEE	1	-0.77117	0.02593	-29.74	<0.0001
OTHEREE	1	-0.90925	0.09419	-9.65	<0.0001
HDD	1	11.39741	1.42866	7.98	<0.0001

Figure 2. Comparison of Actual and Predicted Values for Daily Consumption – Fixed-Effects

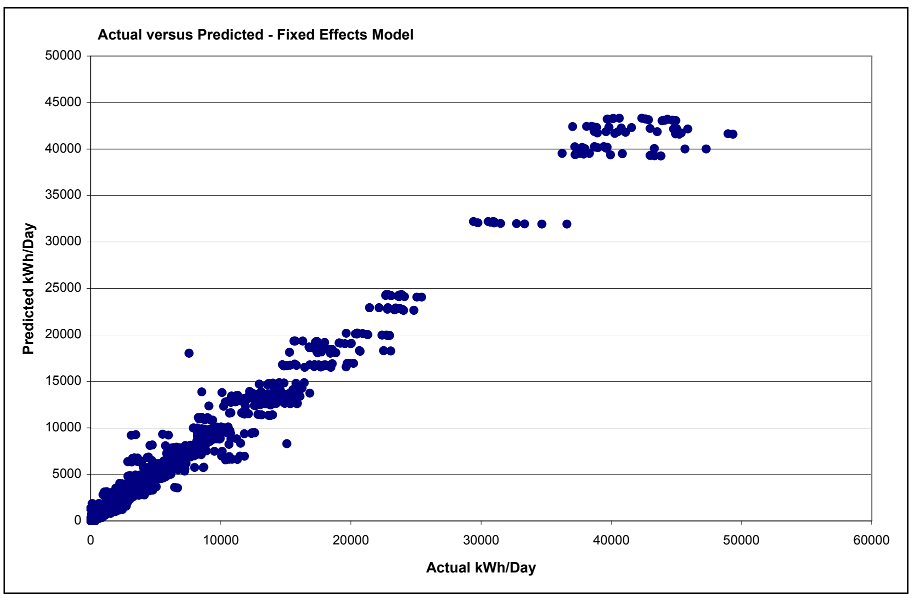
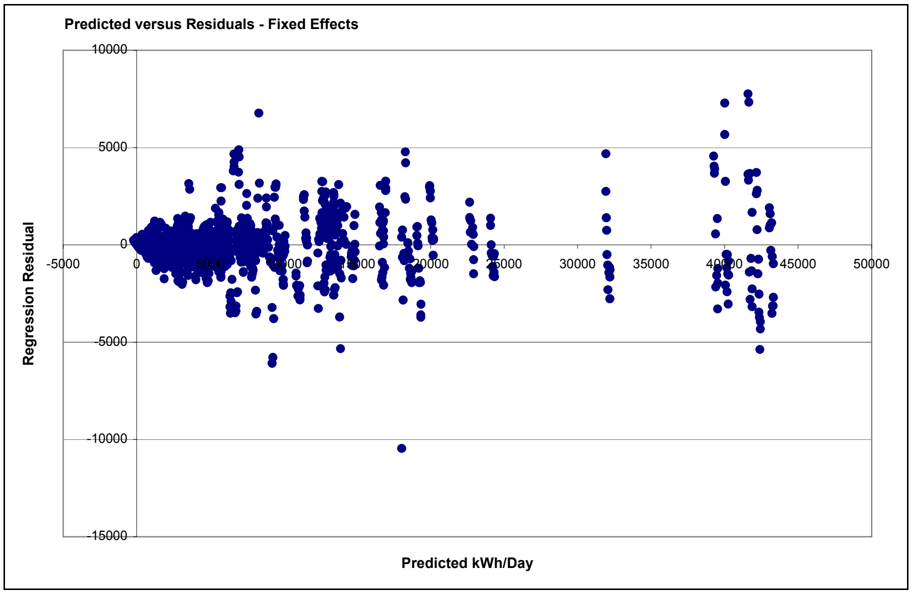


Figure 3. Comparison of Predicted Values and Residuals for Daily Consumption – Fixed-Effects



These analyses results have shown the use of OLS estimator is inappropriate for estimating the parameters of savings models in the large commercial sector. Application of the OLS estimator resulted in coefficients with large standard errors, incorrect signs for all coefficients, and serious problems with heteroscedasticity. The GLS estimators, both in the form of fixed-effects or deviation, significantly improved the statistical properties of the estimated parameters in terms of direction of effect and efficiency. Both visual inspection and application of the White test indicated some degree of heteroscedasticity remained. However, the GLS was shown to significantly reduce heteroscedasticity and result in more robust parameters.

Endnotes

1. First, we estimate model 1. From the model estimation, we output the residuals. Then the squared residuals are run as a dependent variable against an equation that includes all Xs, X^2s, and the product of each pair of Xs. In effect, this involves running the following model:

(1b) $\varepsilon_{it}^2 = \lambda + \delta_1 LIGHTINGEE_i + \delta_2 LIGHTINGEE_i^2 + \delta_3 OTHEREE_i + \delta_4 OTHEREE_i^2 + \delta_5 HDD_{it} + \delta_6 HDD_{it}^2 + \delta_7 LIGHTINGEE_i * OTHEREE_i + \delta_8 LIGHTINGEE_i * HDD_{it} + \delta_9 OTHEREE_i * HDD_{it} + \phi_{it}$

According to the White test, the test statistic is NR^2. This is compared to the critical value distributed as $= \chi^2_{p-1,\alpha}$, where p is the number of covariates. In this case the critical value is $\chi 2$ (df = 10-1=9) [the number of degrees of freedom is determined by the number of independent variables in the model including the squared terms and cross product terms] or 16.92 at the upper 5% percent level. NR^2 = 119.20, which is greater than $\chi 2(9) = 16.92$; hence, we can conclude heteroscedasticity is present in the model.

2. First, we estimate model 3. [Note that computationally it would have been very tedious to obtain the White statistic for model 2. The SAS procedure actually failed to complete the computation, because of insufficient memory due to all of the combinations of the separate facility intercepts and the other independent variables.] From the model estimation, we output the residuals. Then the squared

residuals are run as a dependent variable against an equation that includes all Xs, X^2s, and the product of each pair of Xs. In effect, this involves running the following model:

(3b) $\varepsilon_{it}^2 = \delta_1 LIGHTINGEE^*_i + \delta_2 LIGHTINGEE^{*2}_i + \delta_3 OTHEREE^*_i + \delta_4 OTHEREE^{*2}_i + \delta_5 HDD^*_{it} + \delta_6 HDD^{*2}_{it} + \delta_7 LIGHTINGEE^*_i * OTHEREE^*_i + \delta_8 LIGHTINGEE^*_i * HDD^*_{it} + \delta_9 OTHEREE^*_i * HDD^*_{it} + \phi_{it}$

In this case, the critical value is $\chi 2$ (df = 9-1=8) or 15.51 at the upper 5% percent level. The test statistic NR^2 = 34.70 is still greater than the critical value of $\chi 2$ (8) = 15.51; hence, we conclude heteroscedasticity is present in the deviation model.