Evaluation of a Boiler/Furnace Replacement Pilot Program: Multivariate Regression Analysis Versus PRISM

Irene F. Olson, Washington Gas
Robert P. Trost, George Washington University

Washington Gas' Boiler/Furnace Pilot Program for the District of Columbia is a residential gas conservation program. Energy savings from the program are evaluated here using two regression methodologies. One procedure used is the Princeton Scorekeeping Method (PRISM). Another method is a general multivariate regression model where some restrictive assumptions are placed on the model parameters; Ordinary Least Squares (OLS) is then used to estimate the energy savings. The restricted multivariate model can be easily estimated with most statistical software packages. However, for some applications the implied parameter restrictions may not be justified. Energy savings estimated with PRISM and the multivariate model are compared.

The PRISM analysis indicates that high-efficiency furnaces are fairly impressive energy savers. High-efficiency furnaces show savings of 15.7%; high-efficiency boilers show savings of 7.4%. Less efficient heating equipment does not fare as well. Mid-efficiency furnaces saved only 1.0%; mid-efficiency boilers saved 6.3%. Only the estimates for the high-efficiency furnace participants are statistically different from zero at the 10% level of significance.

The multivariate analysis results parallel those of PRISM. High-efficiency furnaces show energy savings of 15.2%; high-efficiency boilers saved 8.9%. Mid-efficiency heating equipment shows no energy savings; energy usage actually increased after installation. Mid-efficiency furnaces used 8.8% more energy and mid-efficiency boilers used 1.9% more energy. Only the estimates for the high-efficiency furnaces and high-efficiency boilers are statistically different from zero at the 10% level of significance.

Introduction

Washington Gas (WG/Company) is working under District of Columbia Public Service Commission (DC PSC) Order No. 8974 to implement cost effective Demand Side Management (DSM) programs in order to test the options available to reach pre-specified conservation goals. Order No. 8974 established a collaborative Working Group, within a Least Cost Planning framework, to facilitate Washington Gas' development, implementation and testing of programs. The Working Group consists of DC PSC staff, Office of Peoples' Counsel (OPC) staff and D.C. Energy Office (DCEO) staff. This Working Group meets regularly to discuss Company activities and provide input to the development of such activities as program evaluation. The Company's efforts to date reflect this cooperative spirit.

The work contained in this paper has resulted from the implementation of a collaborative process, continues to be discussed and will be a part of Washington Gas' 1992 Least Cost Plan based on further input from the Working Group.

Washington Gas' Boiler/Furnace Pilot Program for the District of Columbia is a residential gas conservation program. Energy savings from the program are evaluated here using two regression methodologies. One procedure used is the Princeton Scorekeeping Method (PRISM). Another method is a general multivariate regression model where some restrictive assumptions are placed on the model parameters; Ordinary Least Squares (OLS) is then used to estimate the energy savings. The restricted multivariate model can be easily estimated with most statistical software packages. However, for some applications the implied parameter restrictions may not be justified. Energy savings estimated with PRISM and the multivariate model are compared.

The Boiler/Furnace Replacement Pilot Program

The Boiler/Furnace Replacement Pilot Program offers a $540 cash incentive for the installation of a high-efficiency

(90%+ AFUE) furnace or a high-efficiency (84%+ AFUE) boiler. Under certain circumstances, a $470 incentive will be paid for a furnace rated 80%-89% AFUE. Low-income customers are eligible for assistance with interest payments incurred in the purchase of the new boiler or furnace.

Collection of Analysis Data for the Boiler/Furnace Program

Consumption data for District of Columbia Boiler/Furnace Replacement Pilot Program participants was collected from Washington Gas billing records. This analysis included all participants who had qualifying heating equipment installed on or before December 1, 1990. A "pre" period was defined for each participant as the twelve months prior to the month of equipment installation; the "post" period was defined as the twelve months after the installation month. The actual month of installation was deleted from the consumption data.

Not all pilot program participants provided consumption data satisfactory for analysis. Some participants either were not gas customers prior to program participation or had non-heating service only. These participants were excluded from the analysis.

Participants in the Boiler/Furnace pilot program were divided into low and non-low income categories. A participant was designated as low income if certified by the District of Columbia Energy Office to be eligible for the Low Income Heating Energy Assistance Program. All participants not eligible for assistance were designated as non-low income.

Using PRISM to Estimate Energy Savings of Heating Customers

A simple comparison approach is often used when estimating energy savings achieved by DSM conservation programs. This comparison can take one of three forms. First, compare DSM program participants' energy consumption $y_p(\text{pre})$ before joining the program to their consumption $y_p(\text{post})$ after joining the program. In this case, energy savings are defined as

$$\text{Savings} = y_p(\text{pre}) - y_p(\text{post}).$$

Second, compare participants' energy consumption in the post period to a control group of non-participants' consumption y_{np} during the same post period. Here, savings are defined as

$$\text{Savings} = y_p(\text{pre}) - y_{np}(\text{post}).$$

Third, compare the savings of the participants to the "savings" of the non-participants. Savings are defined as

$$\text{Savings} = (y_p(\text{pre}) - y_p(\text{post})) - (y_{np}(\text{pre}) - y_{np}(\text{post}))$$

Weather-normalized savings for DSM programs can be estimated using the Princeton Scorekeeping Method (PRISM), a regression method developed by the Princeton Center for Energy and Environmental Studies. This technique is described in Fels 1986 and in Goldberg and Fels 1986. PRISM has been used by Goldberg 1986 to estimate heating season savings of houses heated by natural gas; Stram and Fels 1986 estimated energy savings of houses heated and cooled by electricity.

The approach used by PRISM to weather-normalize energy consumption begins with the estimation of an equation of the form

$$y_{mij} = \alpha_{ij} + \beta_{ij} X_{mij}(\tau_{ij}) + \epsilon_{mij}$$

and (1)

$$X_{mij}(\tau_{ij}) = \frac{1}{N_{mij}} \sum_{k=1}^{N_{mij}} (\tau_{ij} - T_{mij,k})_+$$

where

m = a billing month index.
i = a pre and post period index.
j = a customer index.
k = a day index.

A definition of variables and parameters is given in Table 1. The "+" on the term $(\tau_{ij} - T_{mij,k})_+$ indicates that the value of this term be set to zero if the value in parentheses is negative.

PRISM differs from other weather-normalization procedures in that τ_{ij} is treated as a variable rather than a constant such as 65°F (18°C). The parameters α_{ij} and β_{ij} can be estimated from the M monthly observed energy consumptions by customer j in period i, using standard regression techniques, for any assumed value of τ_{ij}.

Table 1. Definition of Variables and Parameters

D_{mij}	= A dummy variable = 1 if the observation is in the post period (i = 2) for month m and customer j, = 0 otherwise.
$I_{m,mij}$	= 1 for customer j and month m, = 0 otherwise.
I_{mij}	= 1 for customer j, = 0 otherwise.
N_{mij}	= Number of days in billing month m for customer j in period i.
$T_{mij,k}$	= The N_{mij} average outdoor daily billing cycle temperatures for customer j in month m in period i.
$X_{mij}(\tau_{mij})$	= Average heating degree days $HDD_{mij}(\tau_{mij})$ for customer j in month m and period i.
X_{mij}^*	= $D_{mij}*X_{mij}$.
y_{mij}	= Average gas demand for customer j in month m and period i.
α_{mij}	= Base level consumption for customer j in month m and period i.
β_{ij}	= Effect of Heating Degree Days (HDD) on gas demand for customer j in period i.
τ_{mij}	= The "reference" temperature for customer j in month m and period i.
$\delta_\beta=\beta_2-\beta_1$	= The Savings per degree day for Boiler/Furnace Replacement pilot programs.
ϵ_{mij}	= Random error term for customer j in month m and period i.

Once the parameters have been estimated, the normalized annual energy consumption NAC_{ij} for customer j and period i can be estimated from

$$NAC_{ij} = 365\hat{\alpha}_{ij}+\hat{\beta}_{ij}H_0(\hat{\tau}_{ij}) \tag{2}$$

where $H_0(\hat{\tau}_{ij})$ is the total number of heating degree days in a typical year relative to the reference temperature $\hat{\tau}_{ij}$. A typical year represents the average weather conditions over a long-term period for a region.

Given values for NAC_{1j} and NAC_{2j} for all j, the savings attributed to a DSM program can be computed using one of the three comparison approaches discussed above.

The nature of the Boiler/Furnace program makes the selection of a control group problematic. An ideal control group would probably consist of gas customers who replaced heating equipment in the absence of a rebate program. However, Washington Gas does not currently have a database of such customers. Therefore, it will not be possible to give a true measure of free ridership here.

A Simple Multivariate Regression Model of Gas Usage

A researcher may be interested in getting preliminary estimates of energy savings for a DSM program. However, in the beginning months of a pilot program there may not be enough monthly observations in the post period to use PRISM. In this case, a simple multivariate regression (MVR) model could provide a useful preliminary estimate. The disadvantage is that some restrictive assumptions on the parameters of the model are necessary, but these restrictions may not be justified for all pilot programs. For this reason it is important to carefully state what restrictive assumptions are being made. One possible set of simplifying restrictive assumptions for a Boiler/Furnace Replacement pilot program are presented below.

A General MVR Model of Gas Usage in the Pre Period

The following describes a model of the monthly gas demanded by individual heating customers during the six primary heating season months of October through March. The model only considers the six heating season months because, unlike PRISM, it assumes a reference temperature of 65°F. By modeling only the months when heating degree days are greater than zero, the biases caused by specifying an incorrect reference temperature and using OLS rather than Non-linear Least Squares will be minimized. Let the monthly gas demand for customer j in a given heating season month m in the **period before joining** a pilot program be given by

$$y_{m1j}=\sum_{m=1}^{6}\sum_{j=1}^{n}\alpha_{m1j}I_{m,m1j} +$$

$$\beta_{m1j}X_{m1j}(\tau_{m1j}) + \epsilon_{m1j}$$

and (3)

$$X_{m1j}(\tau_{m1j}) = \frac{1}{N_{m1j}}\sum_{k=1}^{N_{m1j}}(\tau_{m1j}-T_{m1j,k})_+$$

where all variables and parameters are defined in Table 1 and

m = 1,6 (months);
j = 1,2,3,....,n (customers).

This equation is similar to PRISM except that base usage given by α_{m1j}, the effect of weather on gas usage given by β_{m1j} and the reference temperature given by τ_{m1j} are allowed to be different in each of the six months. This difference in the reference temperature over time would occur if the inside temperature setting of the house varied from month to month, or if the intrinsic gains from occupants, appliances and the sun varied from month to month. For example, during the holiday season of November and December the indoor temperature setting of the house might be higher to accommodate guests and appliance use might increase.

In Equation (3) y_{m1j} is the average daily use in the pre period for customer j in month m, X_{m1j} is the average heating degree days influencing customer j in month m in the pre period, α_{m1j} is the intercept or "base gas usage" for customer j if $X_{m1j}(\tau_{m1j}) = 0$ for month m in the pre period, β_{m1j} is the effect of weather on gas demand in the pre period for customer j and month m, and ϵ_{m1j} is a random error term for customer j and month m in the pre period. The term $X_{m1j}(\tau_{m1j})$ represents heating degree days and is based on the break-even or reference temperature τ_{m1j}. For many applications this reference temperature τ_{m1j} is taken as a constant of 65°F.

Equation (3) is known as a "fixed effects" regression model since the intercept terms α_{m1j} are constant or fixed for a given customer and given month in the pre period. The underlying assumption of this model is that base gas usage represented by the intercept term α_{m1j} for each customer in month m in the pre period depends on unobserved exogenous factors. Since there are 6n intercept terms in Equation (3), the assumption is that base gas usage α_{m1j} is not only different for each of the n customers, but for a given customer is also different in each of the 6 months.

A General MVR Model of Gas Usage in the Post Period

Let the level of monthly gas demand for customer j in a given heating season month m in the **period after joining** a pilot program be given by

$$y_{m2j}=\sum_{m=1}^{6}\sum_{j=1}^{n}\alpha_{m2j}I_{m,m2j} +$$

$$\beta_{m2j}X_{m2j}(\tau_{m2j}) + \epsilon_{m2j}$$

and (4)

$$X_{m2j}(\tau_{m2j}) = \frac{1}{N_{m2j}}\sum_{k=1}^{N_{m2j}}(\tau_{m2j}-T_{m2j,k})_+$$

where all variables are as previously defined and the subscript 2 refers to the post period.

Placing Some Restrictions on the Parameters

Since the parameters of Equations (3) and (4) have month and pre/post period subscripts, even if heating degree days are the same in two different months, gas demand is different for each individual in each month and in each period. There are several reasons why this might occur. For example, the efficiency of appliances may deteriorate over time or the intensity of use might vary from month to month. Of course, this is an over-parameterization of gas demand. Unless something is known about the values of the parameters in Equations (3) and (4), this equation will not be useful for estimating the energy savings from a DSM program. The only alternative is to place some meaningful restrictions on these parameters. One possible set of restrictions are those made by PRISM. These restrictions are: (a) Base usage is constant over time for any given customer and period. (b) The reference temperatures do not vary over time for any given customer and period. (c) The effect of weather on gas demand does not vary over time for any given customer or period. These three restrictions may be written for all m,i,j as

1. $\alpha_{m1j}=\alpha_{1j}$ and $\alpha_{m2j}=\alpha_{2j}$

2. $\tau_{m1j}=\tau_{1j}$ and $\tau_{m2j}=\tau_{2j}$

3. $\beta_{m1j}=\beta_{1j}$ and $\beta_{m2j}=\beta_{2j}$

Putting the above three restrictions on the parameters of Equations (3) and (4) yields the PRISM model given by

$$y_{m1j}=\sum_{j=1}^{n} \alpha_{1j}I_{m1j} +\beta_{1j}X_{m1j}(\tau_{1j}) +\epsilon_{m1j} \tag{3a}$$

and

$$y_{m2j}=\sum_{j=1}^{n} \alpha_{2j}I_{m2j} +\beta_{2j}X_{m2j}(\tau_{2j}) +\epsilon_{m2j} \tag{4a}$$

PRISM estimates equations (3a) and (4a) separately with non-linear least squares. A researcher may conduct a paired-comparison t-test to test for the significance of energy savings between the pre and post period.

A multivariate regression model can be used to provide a rough estimate of DSM savings while PRISM data is being collected. However, this will require some additional restrictions on the parameters of Equations (3) and (4). Consider the following six restrictions: (a) Base gas usage is the same for each customer in the pre and post periods. The replacement of a boiler or furnace should not have any impact on base usage. (b) Base usage is constant over time for each customer. Although this assumption is not necessary for multivariate regression analysis, it greatly reduces the number of parameters to be estimated. (c) The reference temperature is constant for all customers in the pre and post periods at 65°F. (d) The effect of weather on gas usage is constant for all customers in the pre period at β_1. (e) The effect of weather on gas usage is constant for all customers in the post period at $\beta_2 < \beta_1$. (f) Average heating degree days are greater than 0 for all observations since only winter months are used in the estimation. Combining restrictions (a) and (b), these restrictions can be written as

1. $\alpha_{mij}=\alpha_j$ for all m,i,j

2. $\tau_{mij}=65°F$ for all m,i,j

3. $\beta_{m1j}=\beta_1$ for all m,j

4. $\beta_{m2j}=\beta_2<\beta_1$ for all m,j

5. $X_{mij}(\tau_{ij}) > 0$ for all m,i,j.

Placing these five restrictions on the parameters of Equations (3) and (4) yields the following simple multiple regression model

$$y_{m1j}=\sum_{j=1}^{n} \alpha_j I_{m1j} +\beta_1 X_{m1j}(65°F) +\epsilon_{m1j} \tag{3b}$$

and

$$y_{m2j}=\sum_{j=1}^{n} \alpha_j I_{m2j} +\beta_2 X_{m2j}(65°F) +\epsilon_{m2j} \tag{4b}$$

where all variables are as previously defined.

A Simple MVR Model to Estimate Savings

The ultimate goal of any pilot program evaluation is to estimate the energy savings attributed to the program. These savings can be estimated by combining heating season equations (3b) and (4b) into one simple multivariate regression model of the form

$$y_{mij} = \sum_{j=1}^{n} \alpha_j I_{mij} + \beta_1 X_{mij}(65°F) + \tag{5}$$

$$\delta_\beta X^*_{mij} + \epsilon_{mij}$$

where $\delta_\beta=\beta_2-\beta_1$ and all other variables are defined in Table 1.

Ordinary least squares (OLS) will yield a consistent and efficient estimate of the main parameter of interest δ_β. The savings per degree day from the pilot program are measured as δ_β. The expected sign of this coefficient is negative. If six months of winter data in both the pre and post periods are used to estimate the model, the t-statistic for the coefficient estimate δ_β from the usual regression package output will have 11n-2 degrees of freedom and will test the null hypothesis

$$H_0: \quad \delta_\beta \geq 0$$

against the alternative

$$H_1: \quad \delta_\beta < 0.$$

Estimated Savings for the Boiler/ Furnace Pilot Program

There were 67 participants in the Boiler/Furnace Replacement pilot program with sufficient data for evaluation at the time this work was initiated. Of these participants, 15 purchased high-efficiency furnaces, 24 purchased high-efficiency boilers, three purchased mid-efficiency furnaces and 25 purchased mid-efficiency boilers. Since energy savings depend directly on the efficiency of the new boiler or furnace relative to the one that was replaced, savings were estimated separately for each of these participant groups. Energy savings were estimated with PRISM and with the multivariate regression given by Equation (5).

The mean energy savings from PRISM were calculated from

$$\bar{d} = \frac{1}{n}\sum_{j=1}^{n} (NAC_{1j} - NAC_{2j}) \tag{6}$$

A paired-comparison t-test was used to see if these mean savings were statistically different from zero. Table 2 presents the mean and percent energy savings estimated with PRISM for each type of participant. Percent savings, given by heating unit type and efficiency level, are defined as average energy savings divided by average NAC in the year prior to program participation. Average preparticipation NAC was estimated to be 1206 therms per year for purchasers of high-efficiency furnaces, 1179 therms per year for mid-efficiency furnaces, 1366 therms per year for purchasers of high-efficiency boilers and 1261 therms per year for purchasers of mid-efficiency boilers.

The results from the PRISM analysis of Boiler/Furnace participants presented in Table 2 indicate that high-efficiency furnaces are fairly impressive energy savers. High-efficiency furnaces show savings of 15.7%; high-efficiency boilers show savings of 7.4%. Less efficient heating equipment, however, does not fare as well. Mid-efficiency furnaces show savings of only 1.0%; mid-efficiency boilers show savings of 6.3%. Only the estimated energy savings for the high-efficiency furnace participants are statistically different from zero at the 10% level of significance.

Table 3 presents the OLS estimates of equation (5). The coefficient on the variable X^*_{mit} is an estimate of δ_β and represents a measure of the gain in efficiency after replacing a boiler or furnace. Percent savings, given by heating unit type and efficiency level, are defined as $\frac{\hat{\delta}_\beta}{\hat{\beta}_1}$.

The results given in Table 3 indicate that high-efficiency furnaces show energy savings of 15.2% and high-efficiency boilers show savings of 8.9%. Mid-efficiency heating equipment shows no energy savings; energy usage actually **increased** after installation. Mid-efficiency furnaces used 8.8% more energy and mid-efficiency boilers used 1.9% more energy. The estimated energy savings for the high-efficiency furnace and high-efficiency boiler participants are both statistically different from zero at the 10% level of significance. The energy use increases for the mid-efficiency heating equipment are not statistically different from zero at the 10% level of significance.

The study results also indicate that low income participants may have had non-functioning or inadequate heating equipment in place prior to joining the Boiler/Furnace Replacement Program. These participants, when examined individually, often showed very substantial increases in energy usage after installation of the new heating equipment. Washington Gas is currently conducting a "non-saver" study to determine the factors that caused these energy use increases for some Boiler/Furnace participants.

Conclusions

This study provides evidence that high-efficiency boilers and furnaces have the potential to cut heating energy use dramatically. This benefits the Company as well as the customer as measured by a cost-effectiveness test utilizing the results of this study.

Table 2. Paired Sample t-test of Energy Savings Estimated With PRISM (T-Statistics in Parentheses)

Category	Number of Houses	Mean Savings	Pre Period NAC	Percent Savings
High Efficiency Furnace	15	189.3 (-4.5)	1206	15.7 %
High Efficiency Boiler	22	100.9 (-1.0)	1366	7.4 %
Mid Efficiency Furnace	3	11.5 (-0.1)	1179	1.0 %
Mid Efficiency Boiler	23	79.3 (-1.5)	1261	6.3 %

NOTES: Mean Savings are calculated using Equation (6).
An observation was deleted if any of these conditions held:
(1) β_{ij} less than zero;
(2) Std error of NAC_{ij} greater than $(0.5 * NAC_{ij})$
(3) Std error of τ_{ij} equal to -9.

Table 3. Multivariate Regression Estimates of Energy Savings (T-Statistics in Parentheses)

Category	No. of Houses	Coefficient of X	Coefficient of X^*	RSQUARE	Percent Savings
High Efficiency Furnace	15	0.224 (17.2)	-0.034 (-3.2)	0.96	15.2 %
High Efficiency Boiler	24	0.247 (20.3)	-0.022 (-2.2)	0.96	8.9 %
Mid Efficiency Furnace	3	0.205 (10.2)	0.018 (1.1)	0.98	-8.8 %
Mid Efficiency Boiler	25	0.257 (20.3)	0.005 (0.5)	0.95	-1.9 %

NOTE: Percent Savings are calculated from $\hat{\delta}_{\beta}/\hat{\beta}_1$.
No observations were deleted.

Disclaimer and Acknowledgments

The authors would like to indicate that this paper only reflects their views and opinions. The content of this paper is in no way reflective of the views and opinions held by Washington Gas or by the Public Service Commission of the District of Columbia. This paper discusses work that is ongoing at Washington Gas, District of Columbia Division. At this time, Washington Gas is in a pilot program stage and has not yet fully evaluated all of the pilot programs currently being offered to consumers in the