Energy Trust of Oregon

MEMO:

Date: September 22, 2011
Prepared by: Brien Sipe, Evaluation Project Manager
Subject: Compilation of tankless gas water heater impact studies

Background

This memo includes findings from four impact evaluation results for tankless gas water heaters. Impact evaluations estimate actual savings from installations in the field. Three studies were conducted on the 2007-08 installations of tankless gas water heaters. Evaluation contractors for these studies were: Michael Blasnik and Associates, Demand Research and Stellar Processes. The fourth was conducted by Energy Trust evaluation staff based on data from the 2009 existing homes gas program. Findings have been consistent across studies and provide a high degree of confidence that tankless waters are not cost effective from a societal perspective.

Below is an excerpt from the 2009 Energy Trust existing homes gas impact evaluation on tankless water heaters as well as the memos from the three evaluation contractors in their entirety

Key Findings

- All studies relied on pre/post bill comparisons for large populations of homes with the only measure being installed was a tankless water heater.
- Billing analysis estimates of savings over three years have averaged 65 therms, well short of the original engineering estimate of over 100 annual therms.
    - Original engineering estimates were based on water heating loads well above the average Oregon gas customer's usage.
- All studies filter out homes where fuel switching from electric to gas water heating occurred. This was identified as a frequent occurrence with this particular piece of equipment.
- Energy Trust's planning department will continue to monitor the market for promising high efficiency domestic hot water heating technologies (e.g. condensing tankless water heaters).

Energy Trust 2009 gas impact evaluation: Tankless gas water heaters

Tankless gas water heaters were last evaluated by several contractors in 2009, who examined energy savings during the 2006-2007 program years. Savings from the studies averaged 65 therms, significantly less than the original engineering estimate of 102 therms.

Evidence of significant numbers of homes fuel switching was evident in the 2009 sample, as with the previous studies. Keeping in line with previous studies, a cut-off of 80 therms was used as the minimum threshold for estimated 'baseload' usage when examining savings for the tankless water heaters. This threshold was identified as a balance between causing attrition to the sample and a reasonable floor to indicate whether a home actually had gas water heating prior to the tankless water heater installation.

A difference in differences approach is used to examine both the total loads in the participant and comparison sites, as well as an estimated baseload comparison. Estimated savings, presented below in
Table **1**, averaged 58-66 therms depending on the approach used. Both estimates are comparable to the previous studies which found annual therm savings in the 55-70 range. These new findings are consistent with previous results, and provide more confidence that tankless water heater savings are unlikely to change, and that the measure is simply not cost effective. As with any new product, incremental costs were expected to fall in the years following the introduction of the offering, which has not occurred. Currently, the program is shifting resources to stimulate demand and effect stocking practices of .67 energy factor (EF) tanked gas water heaters (.62 EF products are currently incented).

Table 1 Tankless gas water heating annual therm savings

Load comparison	Participant N	Participant pre use	Comparison pre use	Savings net of comparison (therms)	±95%	Expected savings (therms)
Tankless baseload only	230	229	228	**66**	13	65
Tankless total load comparison	230	776	764	**58**	29	65

Billing Analysis of Tankless Gas Water Heater Rebate Program

Michael Blasnik (Michael Blasnik & Associates)

Draft #2 January 8, 2009

The objective of this project is to compare and contrast alternative approaches to billing data analysis. In this particular instance, the billing analysis is intended to assess the energy savings from a tankless gas water heater rebate program

Data Collection and Analysis Approach

ETO provided tracking system data and monthly gas billing data for customers who participated in the tankless gas water heater rebate program. ETO also provided daily outdoor temperatures from January 1, 2000 through August 31, 2008 for 11 weather stations.

The tracking data included 1003 measure level records for 622 customers that received rebates between July 2006 and September 2008. There were 381 records for other measures also rebated at these same premises including 67 clothes washers, 170 gas furnaces, and 68 insulation upgrades. Overall, there were 368 customers that only participated in the tankless rebate program and 254 customers that also participated in other programs. The data included the floor area for each home, the type of measure and installation date for each measure, and also included the brand, Energy Factor, and installed cost for the tankless water heater.

The billing data included 22,696 meter readings for the 622 participants spanning from October 2005 through September 2008. Billing data were also provided by ETO for a comparison group of 28,723 customers randomly sampled from the same zip codes as participants . This file included 1,153,642 meter readings. ETO staff “cleaned up” the billing data by combining any estimated readings into the subsequent actual readings. The billing data also included weather station assignments for each customer.

The billing data were separated into pre and post retrofit periods based on the starting and ending measure installation dates for each participant. For the comparison group, pseudo treatment dates were randomly assigned from the pool of pre-2008 participant treatment dates. Pre and post treatment meter readings were eliminated if they occurred more than 14 meter readings and 450 days before or after the treatment date. In addition, comparison group meter readings were excluded from the analysis if they occurred outside of the range of meter reading dates found in the participants’ data.

The pre and post treatment billing data for each participant and comparison group customer was weather-normalized using a variable-base degree day regression model similar to PRISM. This model differs from PRISM in that it employs a Bayesian approach to estimating the balance point temperature which helps to avoid extreme balance point temperature estimates.

Weather normalization results were screened for reliability by removing cases where:

- the regression model fit was not very good (R-squared <0.7 or CVnac>20%)
- there was insufficient data (<180 days or <40% of a normal year's HDD or max HDD/day- min HDD/day < average HDD/day)
- the baseload usage estimate was negative or the relative standard error of the baseload usage estimate was greater than 100% of the estimate (and also greater than 40 therms, to allow for low baseload use)
- the change in total usage was greater than 65%
- comparison group cases with total or baseload gas usage outside the range of usage found among participants
- the building was not listed as single family.

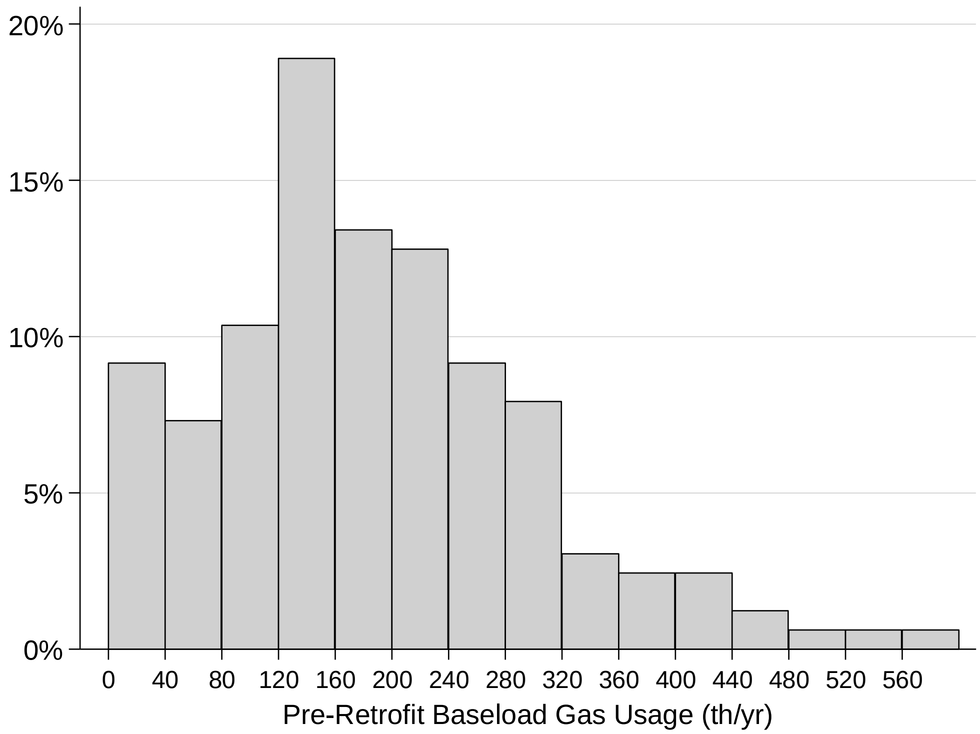
The data screening resulted in 251 participants and 10,164 comparison group cases with apparently reliable weather adjusted usage results. For participants, 40% of cases passed the screening, 38% were eliminated due to lack of data, 21% were eliminated due to apparently unreliable usage results (bad fit or negative or uncertain baseload) and 1% were eliminated as outliers. The relatively high attrition rate is related to the lack of post-treatment data for many cases -- just 216 of the 622 participants were treated more than a year prior to the end of the available usage data.

The net savings analysis involved calculating the mean savings for the participants and then subtracting the mean savings of a weighted comparison group. Comparison group cases were weighted to match the participant group using a post-stratification on weather station (3 stations), pre-treatment annualized total gas usage (7 bins) and pre-treatment annualized baseload gas usage (typically 6 bins). This weighted matching method provides a flexible way to improve the comparability of the comparison group without requiring many of the assumptions inherent in a regression-based approach.

Findings

The initial analysis found net annual savings of 44 therms in total gas usage (54 therms in baseload savings and -10 therms heating savings). This level of savings is considerably smaller than the 102 therm working savings estimate. More notable was the average pre-participation baseload gas usage which averaged just 188 therms -- less than the 200-250 therms expected in homes with gas water heating. This lower level of usage might be consistent with apartments or low occupancy homes with small hot water loads.

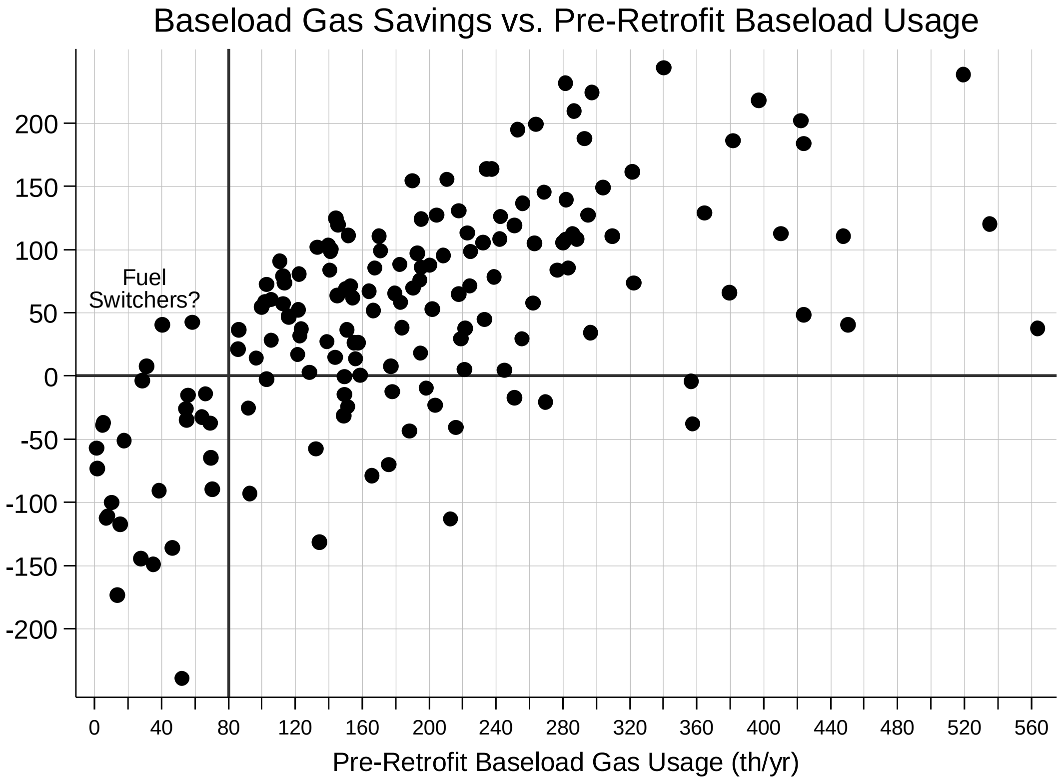
The figure on the next page is a histogram of baseload gas usage for the 164 participants with reliable analysis results and that only received a tankless water heater rebate.



The figure shows that a significant fraction of cases had pre-treatment baseload usage that appears to be too small for supporting a conventional gas water heater. Nearly 10% of the participants had baseload gas usage of less than 40 th/yr and another 7% had baseload usage between 40 and 80 th/yr. Standby losses for a conventional tank gas water heater can be expected to total about 65 th/yr making it highly unlikely that a customer with a standard gas water heater has a baseload usage much less than that. Baseload usage of 80 th/yr may be feasible for a household with very low hot water use (perhaps a 1 person household), but lower usage rates make it less and less likely that a conventional gas water heater was used. There are three reasonable explanations for the low baseload usage estimates in some homes:

- the customer may have had electric water heating and switched to gas when purchasing the tankless unit;
- the customer may have already had a tankless unit and also used relatively little hot water leading to low baseload usage; or,
- the low baseload usage is due to an anomaly in the billing data or weather normalization analysis such as a long summer vacation or a meter reading error.

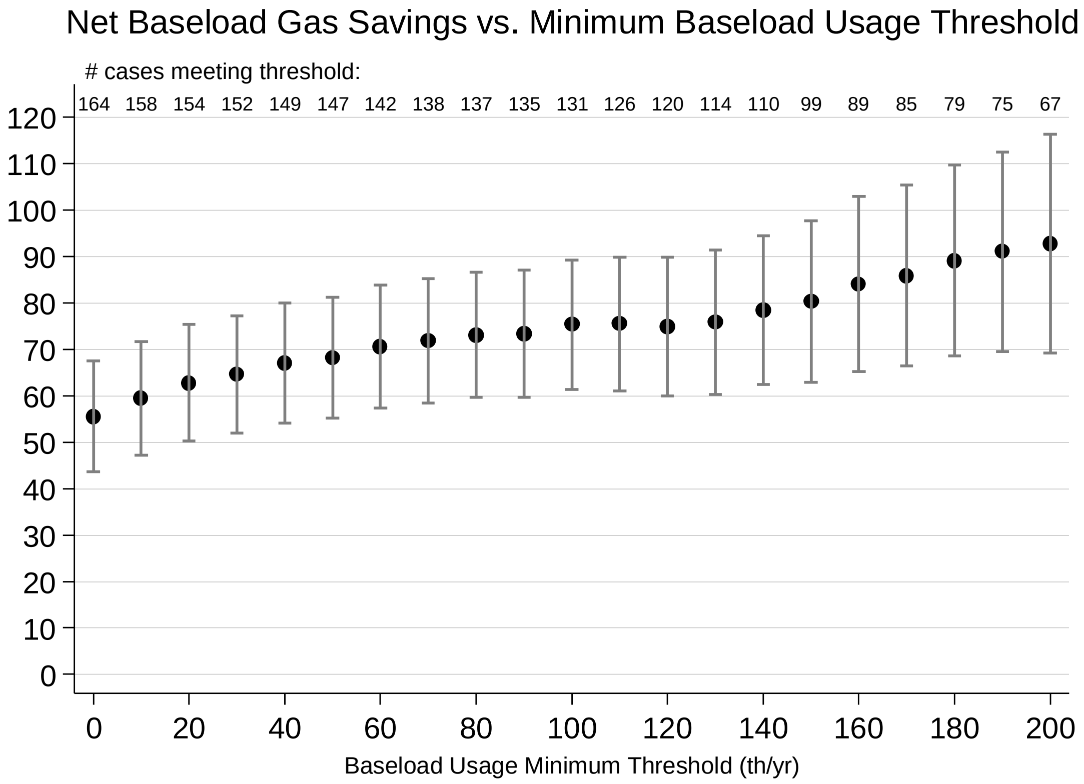
The following graph shows baseload gas savings compared to pre-treatment baseload usage.



It appears that many customers with low pre-retrofit baseload usage experienced a large increase in baseload use. The vertical line shows a baseload of 80 th/yr and most customers with usage below that level had increased baseload use after the water heating retrofit. Based on this analysis, a request was made to ETO for the electric billing data of these customers to see if a large decrease in electric usage was evident. ETO provided simple average daily electric usage for the years before and after retrofit. The electric data showed that more than 90% of the customers with baseload gas use below 80 th/yr experienced a reduction in electric usage of more than 1000 kWh/yr and about half showed a reduction of more than 3000 kWh/yr. These findings support the hypothesis that many of these low gas baseload use customers switched from electric to gas water heating.

Ideally, a quick customer phone survey could properly identify fuel switchers and an extra question on the rebate form could address this question in the future. But time and budget constraints precluded that approach for this analysis.

In lieu of direct fuel switching data, we explored creating a minimum baseload usage threshold for the analysis to estimate the actual gas savings from tankless gas water heaters excluding fuel switching cases (where the analysis would need to also analyze electric usage changes more closely). There is no clear line of demarcation between very low gas water heating loads and fuel switchers, although some value between 40 therms and 120 therms is likely. If the threshold is set too low, then mistakenly included fuel switchers may lead to understated savings but if the threshold is too high then the savings may be overstated as low users are eliminated. A sensitivity analysis was used to help identify a threshold and assess its impacts. The figure below shows the estimated overall net baseload savings using a wide range of minimum baseload usage thresholds.



The leftmost point with capped line shows the net baseload savings and 90% confidence interval for all cases with baseload usage greater than 0 therms (i.e., all cases). The next point with confidence interval shows the net savings after eliminating all cases with pre-treatment baseload usage less than 10 therms. The third point shows the net savings with a 20 therm minimum baseload and so on. The savings increase fairly steadily until about 80 therms and then level out before increasing again above 130 therms. The numbers along the top of the graph show the number of participants that meet the threshold. At 80 therms, there are still 137 of the 164 participants but at 200 therms, just 67 participant remain. Based on engineering judgment and this graph, a threshold of 80 therms was selected as the best trade-off between bias from fuel switchers vs. bias from omitting low hot water loads. The graph also illustrates how the net savings would be relatively unaffected by shifts in this threshold anywhere from 60 to 140 therms. The net savings analysis was performed based only on customers with a pre-treatment annual baseload usage of at least 80 therms.

The net gas savings for participants that only received the tankless water heater rebate are summarized in the table on the next page.

Table 1. Net Gas Savings Analysis Results (therms/year)

	Pre	Post	Save	Net Savings	
Participants (n=137)					
Baseload	218	147	71	73 (±13)	33.5% (±6.2%)
Heating	489	480	9	-5 (±16)	-1.1% (±3.3%)
Total	707	627	80	**68 (±16)**	9.6% (±2.3%)
Comparison Group (n=6511)					
Baseload	218	221	-3		
Heating	476	461	14		
Total	694	682	12		

The 137 participants reduced their annual gas usage by an average of 80 therms – from 707 to 627 therms. The comparison group experienced an average 12 therm reduction in usage leading to net annual savings of 68 therms which is equal to almost 10% of annual usage. The 90% confidence interval on the net savings spans from 52 to 84 therms. The ETO working savings assumption was 102 therms and so this finding would indicate a savings realization rate of 67%.

The 68 therms net savings is composed of a net 73 therm decrease in baseload gas usage and a 5 therm increase in heating usage. This heating usage increase is small enough that it could be random noise although a small increase would be expected to the extent that water heater tank standby losses contribute some useful heat in the winter.

The total net savings of 68 therms equals 31% of the 218 therm average pre-treatment baseload usage. Cooking, clothes drying, and other gas baseload uses are unknown for these homes, but might average about 15 th/yr. Using this assumption, the water heating load savings may be about 33% (68/203). This percentage savings would be consistent with a reasonable pre-retrofit Energy Factor of about 0.54 given the rated 0.81 average Energy Factor of the tankless units -- (0.81-0.54)/0.81 = 33%.. The ETO working savings estimate of 102 therms would be consistent with some combination of a larger hot water load and a lower existing Energy Factor. For example, a hot water load of 267 therms and an existing Energy factor of 0.50 would produce estimated savings of 102 therms.

The gas savings results appear to be generally consistent with the observed baseload gas usage and rated/estimated Energy Factors of the pre and post retrofit equipment. But this finding may appear somewhat at odds with recent research in California that indicated that tankless gas water heaters perform worse than their rated Energy Factor under typical household hot water usage patterns due to cold heat exchanger start-ups associated with small hot water draws. The Energy Factor test only uses a few large hot water draws. The California research suggests that the real world Energy Factor may be closer to 0.72 for a tankless unit rated at 0.81. However, tankless water heater manufacturers have pointed out that a conventional tank unit will perform below the rated Energy Factor if hot water loads are lower than the test values as the standby losses become a larger proportion of the usage. Given the relatively modest estimated hot water usage of 200 therms (vs. about 250 therms expected from the Energy Factor test), a basic engineering analysis suggests that a conventional gas water heater would have an effective Energy Factor about .06 below the rated value. The measured savings of 68 therms are consistent with making both of these Energy Factor adjustments -- a tankless unit at EF=0.72 and a conventional unit at EF=0.48 with a pre-retrofit usage of 203 therms yields estimated savings of 68 therms. This analysis provides further support for the billing analysis finding.

In addition to the analysis of the 137 participants who only received the tankless rebate, we analyzed the usage of an additional 79 participants who received other measures (including 66 participants who received heating-related measures) . The net baseload gas savings for these 79 participants was 72 therms -- essentially identical to the 73 therms in baseload savings found for the tankless-only participants. This finding lends further support to the primary analysis results.

Ideally, a larger scale billing analysis would be performed to follow up on these findings using larger samples and supplemented with a survey about water heating fuels.

To: Phil Degens, Brien Sipe

From: Marvin Horowitz

re: Final Findings of the Tankless Water Heater Program (unbalanced panel billing analysis)

Here is a description of the tankless water heater program data, analysis, and findings:

- I used only those accounts that received tankless water heaters (n=293).

- I used all the available meter reading periods for each participant and nonparticipant, the maximum number being 35, and none having less than 10.

- I used actual values of average daily hdd and readings, per month as model variables. I did not remove or modify 0 values. However, I removed from all models a small number of periods in which average daily readings were equal to or greater than 30 therms per day per month.

- To capture average program impacts I used a dummy variable where 0 was the value for the periods prior to the tankless water heater installation and 1 was the value for periods after installation

- Separate but identical models were estimated for datasets containing participants alone and participants combined with nonparticipants. Also, models were estimated for datasets containing all average heating degree days per period, those with average heating degree days of less than 10, and those with heating degree days of less than 3. All six models were estimated using panel least squares with variables in linear form and with a White coefficient covariance correction for the standard errors.

- I calculated total annual therm savings by multiplying the coefficient of the impact variable by the number of program participants and then the number of days in an average year, i.e., 364.25.

- Several versions of the billing analysis were performed, and the results of all of these analyses are available on request. After experimenting with many different model and data specifications, such as using a log-log specification, the final models chosen for this study appears to be the most reliable and direct for calculating program savings. Due to the homogeneity in cross section observations, as shown in Table 1, and the limited number of end uses for natural gas, cross section fixed effects and weighted least squares models were eliminated from final consideration.

Table 1 contains descriptive statistics for program participants and nonparticipants for the key variable, average daily therms per period. These are provided based on the three different heating degree day based samples; that is, the entire set of periods, only those periods whose average daily heating degree days were less than 10, and only those periods whose average daily heating degree days were less than 3.

Table 1: Descriptive Statistics for Average Therms per Period (AVGREADING)

	Average Therms per Period		
	All Periods	AVGHDD<10	AVGHDD<3
PARTICIPANTS			
Mean	1.99	0.92	0.59
Median	1.45	0.69	0.48
Maximum	28.09	12.93	7.45
Minimum	0.01	0.01	0.01
Std. Dev.	1.84	0.84	0.48
Observations	9164	4583	2612
NONPART.			
Mean	2.18	1.19	0.86
Median	1.48	0.81	0.62
Maximum	29.97	29.10	29.10
Minimum	0.02	0.02	0.02
Std. Dev.	2.24	1.45	1.13
Observations	45078	24125	14031

Table 2 contains the estimated savings for program participants; all models and their associated statistics are contained in the Excel spreadsheet "HORO_TANKLESS-FINALPROGRAMSAVINGS.xls." One important finding is that the savings estimates increase as the number of average daily heating degree days per period declines. Another important finding is that estimated savings more than double when nonparticipants are added to the models.

Table 2: Summary of Tankless Water Heater Savings

IMPACTS (Therms)	Participants, Only	w/Nonparticipants
customer n	293	1893
	All Periods	
Mean Daily Impact	-0.107	-0.329
Total Annual Impact	-11,421	-35,087
Realization Rate	38%	117%
	AVGHDD<10	
Mean Daily Impact	-0.132	-0.343
Total Annual Impact	-14,136	-36,592
	47%	122%
	AVGHDD<3	
Mean Daily Impact	-0.165	-0.365
Total Annual Impact	-17,596	-38,978
Total Annual Impact	59%	130%
Expected Annual Impact	29,886	

Expected savings of approximately 30 thousand therms is calculated by multiplying the ex ante estimate of 102 therms of savings per tankless water heater by the 293 participants. Depending on the sample used for the model, the realization rate varies from a low of 38 percent to a high of 130 percent.

Tankless Water Heater Program Impact Evaluation

For

Energy Trust of Oregon

By

David Robison, P. E.

Stellar Processes

Stellar Processes, Inc.

1033 SW Yamhill Suite 405

Portland, OR 97205

(503) 827-8336

September 30, 2011

Table of Contents

List of Tables

List of Figures

Executive Summary

This report analyzes the energy saving due to the installation of a tankless water heater to replace a conventional gas water heater. These installations were conducted under the program of the Energy Trust of Oregon.

This report is part of a series conducted by a panel of evaluators in order to assess and compare different analysis applications. The source data were prepared by Energy Trust staff and the total program participation numbers were not provided. However, for the observations provided, one draws the following conclusions:

- Net savings averaged 71.1 ± 23.3 annual therms
- Review of a comparison group showed no significant impact due to background conditions
- Participant behavior was very highly variable which interfered with the analysis. Operational changes often showed up as changes in space heating. However, a Conditional Demand Model (CDM) did not find statistically significant impacts for other conservation measures conducted at the same time.

Energy Usage Analysis

Documentation Of Analysis Methodology

Temperature Regression Model

Studies of this sort frequently use the PRInceton Scorekeeping Method (PRISM)[1] to weather normalize the consumption data and to remove 'noise' due to the weather which could influence consumption before or after treatment. The difference between consumption before and after treatment represents the energy savings due to the program.

[1] Since PRISM is a trademarked term and the proprietary software was not used, the models used here are best described as PRISM-like.

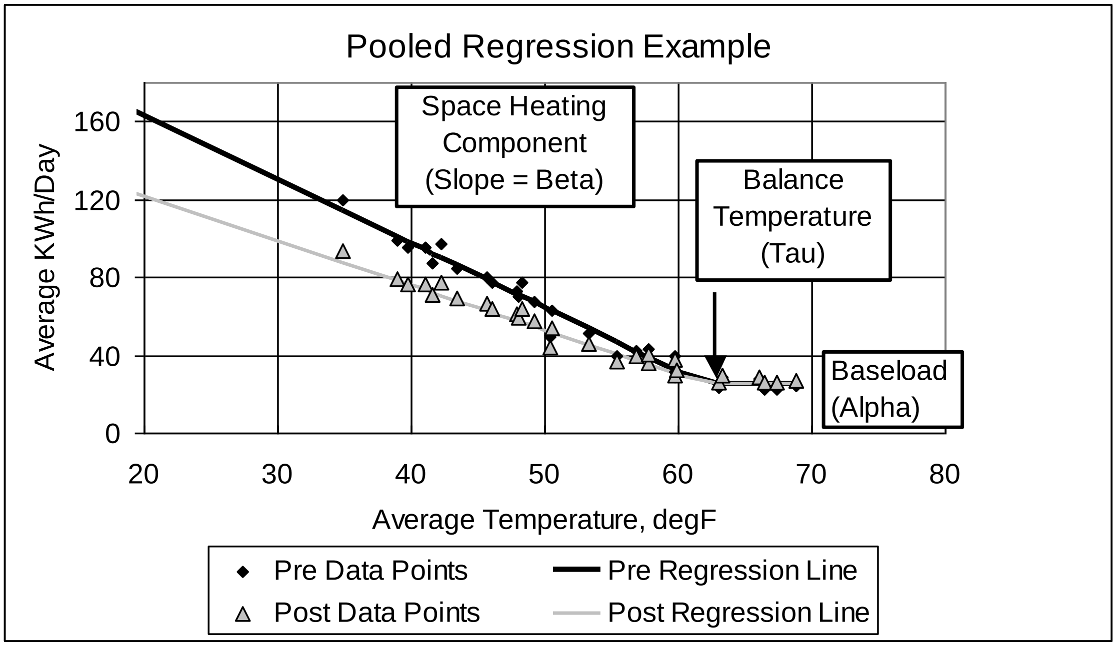


Figure 1 Temperature Regression Example

PRISM-like models apply regression to separate the electric bill into a baseload and seasonal, weather-dependent component. This output is used in the analysis to determine whether there have been changes in the baseload or weather dependent component after treatment. The weather dependent component is assumed to represent primarily space heating. The slope of this space heating with respect to temperature is referred to as “beta”. Space heating occurs only below a balance temperature, referred to as “tau”, which is unique to each home. The balance temperature depends on the thermal integrity of the house, the preferred thermostat setting of the customer and other behavioral factors. During the summer, when there is no space heating, consumption includes only a baseload component, referred to as “alpha”. In the illustration, data observations from both pre and post-retrofit are pooled in the regression model in order to develop coefficients that represent the change due to the treatment.

In practice, one tests a series of regression models over a range of balance temperatures and chooses the one with the best fir to the observations. In this process, one is testing for the variable degree-day description that best fits the individual participant’s energy consumption behavior. Thus, one creates an independent variable that represents the heating degree-days (HDD) defined as the average of (Balance Temperature - Average Daily Temperature) for the metered interval.

Normalizing Results

To better visualize the change in consumption due to the program, one would prefer to look at results normalized across any confounding variables. To normalize for weather, one computes predicted consumption from the appropriate regression model evaluated at weather conditions that correspond to ling-term average conditions. The sum of both baseload and weather-dependent components provides the Normal Annual Consumption (NAC). That is, NAC is the typical total annual energy consumption during a “normal” weather year.

Daily temperature data provided by National Oceanic and Atmospheric Administration (NOAA) were used to build the temperature file used by the regression. To normalize weather to long-term average, one often uses long-term weather parameters also supplied by NOAA. In this case, we had source data that allowed computation of HDD for weather stations in Oregon. Although the available weather data spans about six years, the average HDD are quite similar to those computed from NOAA long-term average monthly observations. Since these averages computed from the daily six-year observations better matched the procedure used in this evaluation, we applied the six-year averages to estimate “normal” long-term NAC.

Removal of Outliers

A frequent problem with regression studies is that a large portion of the study population must be removed due to the inability to form a successful regression model. Such problems might be caused by gaps in the data, vacancies or changes in the occupant’s behavior. In this case, we reviewed the data carefully in order to preserve as many cases as possible.

Programmatic Impact

Programmatic impact on consumption was evaluated using a traditional quasi-experimental design. The design compares the participants to a similar but untreated group. In a true experimental design, members would be assigned randomly to either the treated or the comparison group.[2] This approach is not possible for an actual program where interested customers are allowed to participate. Hence, the design is considered “quasi” experimental.

Ideally, one would draw non-participants from a pool of future program participants to minimize the possibility of any self-selection bias. Regression analysis of the

---

[2] Cook, Thomas and Campbell, Donald, Quasi-Experimentation, Design and Analysis Issues for Field Settings, Houghton Mifflin Co., 1979. Campbell, Donald and Stanley, Julian, Experimentation and Quasi-Experimental Design for Research, Houghton Mifflin Co., 1963.

Comparison Group (control group) followed the same procedures developed from the analysis of the Participant Group.

The analysis uses a standard pre/post cross sectional consumption (billing) analysis. The weather normalized annual consumption (NAC) before the treatment establishes a baseline, which can then be compared to weather normalized consumption after the treatment. The difference in consumption determines gross savings. That is:

Gross savings = NAC(pre) - NAC(post)

Gross savings are determined for the comparison group in the same way. The participant savings are corrected for any consumption change apparent in the comparison group. The result is net savings attributable to the program. This difference of differences approach is traditionally used in DSM evaluation to “net out” savings due only to the treatment.[3] Results are reported in terms of the average savings per dwelling unit.

Sample Disposition

The total number of program participants is not clear. The “Data Description” documentation refers to 622 sites. However, the list of measures “Tankless Measures” shows 698 participants. Of those, billing data were provided for 667 sites. For 161 cases, billing data were missing or insufficient in duration. An analysis mode was not successful for 32 cases, leaving 505 cases for analysis. The resulting set of cases represents 72% of the documented participants, which can be judged to be a representative sample.

Table 2. Sample Disposition

	698	Measures Installed
96%	667	With billing records
-23%	(161)	Insufficient records
-5%	(32)	No useful model
72%	505	Cases for analysis

[3] Fels, M. The Princeton Scorekeeping Method: An Introduction, Princeton University, Center for Energy and Engineering, Princeton, NJ, PU/CEES 163. Fels, M., Special Issue Devoted to Measuring Energy Savings: “The Scorekeeping Approach”, Energy and Building, 9(1-2), Feb/Mar 1986.

Modeling Approach Options

Ideally, one would analyze a group that did not receive any other treatment in order to isolate the impact of the treatment alone. However, that is not practical for many programs since participants often choose as multiple conservation offers. For this program, the original intent was to look separately at participants who received water heating measures and those that also received space heating measures. One would expect that the first group would show consumption changes only in the baseload component while the second group would have changes in both the baseload and space heating components. For the first group one could look at a regression model that pools pre and post-retrofit records with a variable to represent the baseload offset due to treatment. The regression model is then:

$$Average\ consumption = \propto\ + \beta_1 * HDD_n + \beta_2 * D$$

where β_2 is the coefficient associated with a treatment variable and HDD_n are the Heating Degree Days associated with the monthly observations.

The advantage of this model is that it allows all the observations to be pooled together with minimal explanatory variables. Thus, it is suitable even if there are relatively few post-retrofit observations available. The disadvantage is that it requires that the participants did not change consumption in any way that affects space heating.

For the second group, the regression model needs an additional variable to represent the change in space heating. The regression model is of the form:

$$Average\ consumption = \propto\ + \beta_1 * HDD_n + \beta_2 * D + \beta_3 * D * HDD_n$$

where β_3 is the coefficient associated with change to space heating.

Because this modeling is more complex, the second group will require more data records that fully span the heating season. Cases that were installed late in the year often lack sufficient records to establish any change in space heating behavior. Thus, this model is more likely to fail due to insufficient billing records. In the following discussion, we refer to the “water-heating” group and the “space-heating” group, although it should be clear that the second group is actually both space and water-heating.

Analysis Results

Modeling Approach

As it turned out, some of the participants that were supposed to be water-heating only demonstrated such operational changes that the full space heating model was required. Table 2 shows that about half the participants were about equally divided between water-heating and space-heating cases. For each case, we reviewed both types of model to determine which appeared to provide the best fit to observations. That is, we did not rely solely on the stated measure list. Criteria for selection of acceptable models included assessment of the following steps:

1) R^2 of less than .75 indicates a poor model fit, use the better model.
    a. However, R^2 of 1.00 indicates a faulty coefficient due to insufficient data
2) If the t-test for water-heating treatment variable is at least 2, use the water heating model.
3) If neither model has an acceptable R^2, reject case as no useful model.

A further complication arose when it turned out that some of the participants apparently substituted the gas tankless heater for an electric heater. (Switching between fuels was not supposed to be allowed). Since these participants were atypical, they were also removed from the analysis group.

Table 3. Final Modeling Disposition

	Sum of Cases	Useful Model	No Useful Model	Fuel Switch Cases
Water Heating Model	276	198	22	56
Space Heating Model	229	210	10	9
Total Cases Modeled	505	408	32	65
Percent of Billing Records	76%	61%	5%	10%

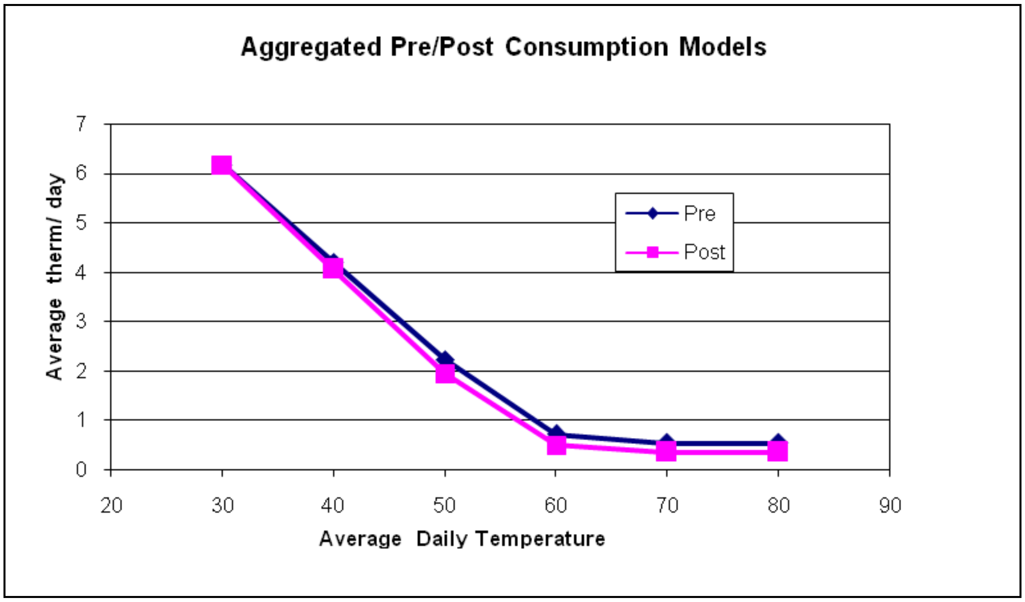
Analysis Results

In general, results proved highly variable which interfered with drawing conclusions. Both the water-heating and the space-heating groups showed very similar average savings. In part, this is due to the higher variability when space heating is involved. Estimated savings from the two types of models are compared in Table 4. Both approaches give results that are statistically significant. As shown in Figure 2, the difference between pre and post consumption can be generally viewed as an offset to the baseload. To normalize for weather, we prepared an aggregate chart of the regression results as a temperature-dependent model of energy consumption pre- and post-retrofit. This aggregation preserves the temperature dependency of the savings but is independent of any specific climate. Note that Figure 2 does not provide sufficient information to evaluate the annual savings. That is because one must also know the number of days that each site experiences a particular average temperature.

Table 4. Modeled Savings Estimates, Annual therms Saved

	Mean	Standard Deviation	Standard Error	Count	95% CL Lower	95% CL Higher
Water Heating Model	71.7	155.3	11.0	198	50.0	93.5
Space Heating Model	71.2	231.0	15.9	210	39.8	102.6
Total Cases Modeled	71.4	197.9	13.8	408	44.3	98.6

Figure 2. Aggregated Pre/ Post Consumption



Distribution by Size Strata

Staff provided a size stratum descriptor for non-participants. Although the definition was not clear, we interested the definitions to be as follows:

1. Less than 800 annual therm
2. Greater than 800 but less than 1100 annual therm
3. Greater than 1100 therm
4. Not present in participant set

The two modeling approaches demonstrated little difference in the distribution by size strata.

Table 5 shows that most of the cases fell into the first stratum. This distribution is later applied in examining the non-participant group.

Table 5. Distribution by Size Stratum

Size Stratum	Water Heating Model	Space Heating Model	Total Cases Modeled
1	62%	59%	60%
2	26%	27%	27%
3	12%	14%	13%

Distribution by Weather Location

Distribution by weather location is not expected to affect the savings estimate because the water heating consumption is part of baseload and is independent of climate. Table 6 shows the distribution of cases modeled by their source weather data.

Table 6. Distribution by Weather Location

Weather Location	Water Heating Model	Space Heating Model	Total Cases Modeled
1 Astoria, Oregon Coast	3%	3%	3%
2 Eugene, Lower Willamette Valley	19%	21%	20%
9 Portland, Upper Willamette Valley	77%	77%	77%

Non-Participant Group

As a check on the methodology, it is useful to examine a comparison or control group. The purpose is to assure that there were no underlying background conditions that might have influenced any observed impact. For example, a sudden rate increase might have induced participants to undertake conservation actions on their own. If so, one would like to "net out" any such impacts. For this check, Energy Trust staff provided records for a non-participant group that installed a similar conservation measure but in subsequent years. Such a group is well suited for comparison -- one hopes that any self-selection bias is eliminated since these cases also chose to participate, just at a later date.

We followed the same analysis procedure for the comparison group, including investigating both the water-heating and space-heating models. **Error! Reference source not found.** and Figure 4 show there is little evidence of consumption change for non-participants. We examined results by the distribution of annual consumption as represented by size stratum. This turns out to be important because some evidence of background conservation was present based upon size stratum. As shown in **Error!**

Reference source not found., there is evidence of negative savings in the small stratum – that is, their consumption increased. Meanwhile, the largest stratum showed evidence of positive savings; their consumption decreased. These results are what would be expected for "reversion to the mean". That is, if there were random changes, cases at the two extreme tails would tend to move toward the midrange in the second year. The group in the mid-range shows no significant savings. Mean results for all size strata are somewhat misleading because they fail to account for the differing distribution of size strata among the participant cases. When the participant weights are applied, the top and bottom strata tend to cancel out resulting in a weighted average that is not significant compared to the expected confidence limit. Based on these results, we assume that there was no significant background to be netted out of gross savings results for either of the two modeling approaches.

Table 7. Non-Participant Group Savings by Size Stratum

Modeling Approach	Size Stratum	Weight in Participant Cases	Mean	95% CL	t-Statistic*	P Value (2-tailed)*	Count
Water Heating Model	1	60%	-17.7	11.6	-2.99	0.00	271
	2	27%	5.6	15.7	0.99	0.32	338
	3	13%	39.5	15.7	4.96	0.00	329
	Weighted Average		-4.7	13.1			
Space Heating Model	1	60%	-14.9	9.9	-2.98	0.00	236
	2	27%	6.0	14.4	1.23	0.22	321
	3	13%	34.0	14.4	4.66	0.00	338
	Weighted Average		-2.4	11.7			

*t-statistic and P value based on difference of means test between NAC1 and NAC2

Figure 3. Space Heating Models, Non Participants

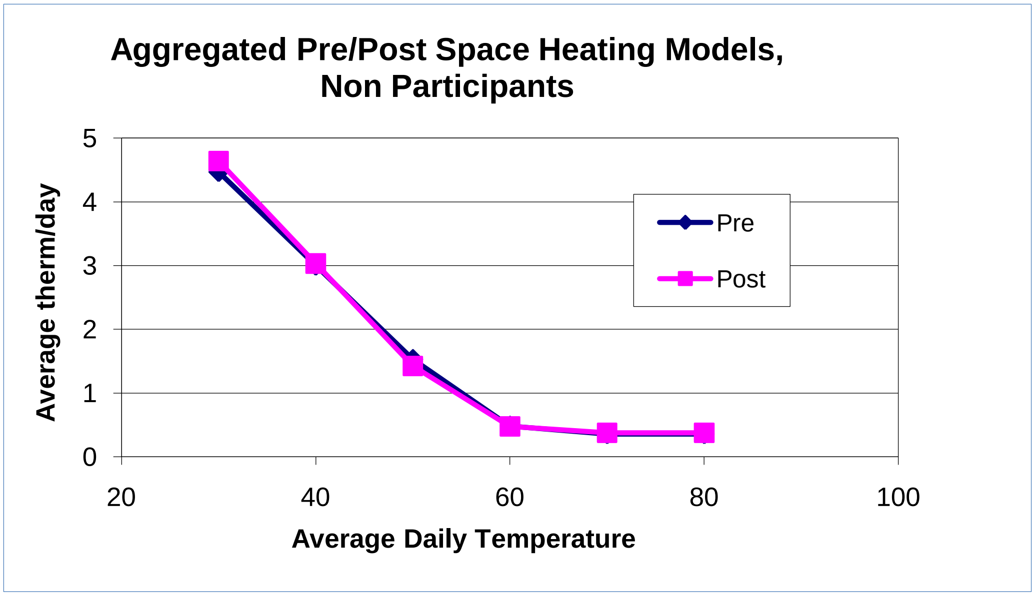
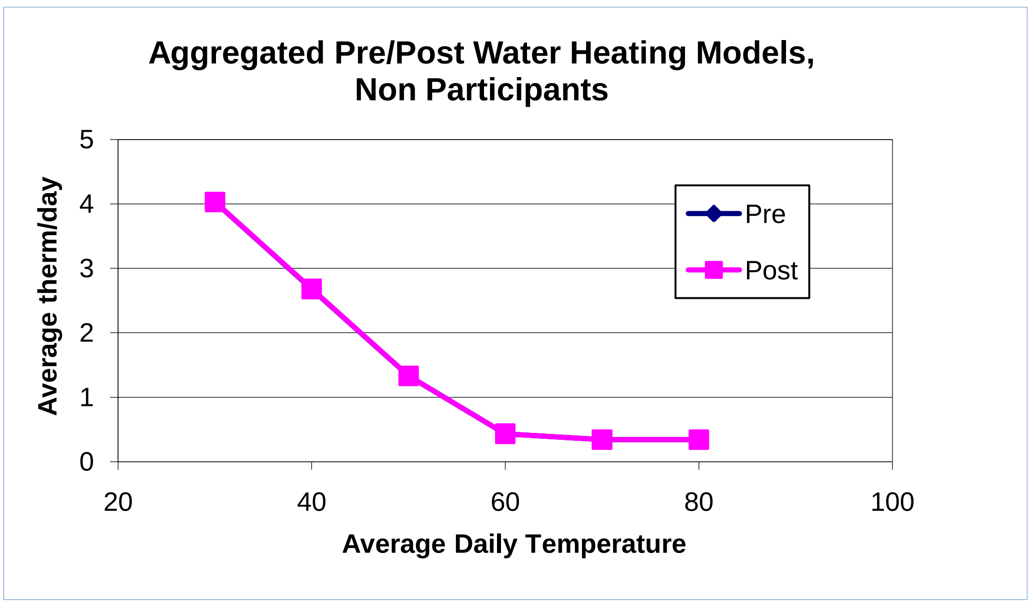


Figure 4. Water Heating Models, Non Participants



Conditional Demand Model

The close simularity of the two groups is puzzling because one would expect that the particpants with space heating changes would have higher savings. Most of the space heating cases installed a high efficiency furnace. The savings for this measure are small and could easily be hidden in the high variability of the results. Meanwhile, many of the water heating cases should have had additional savings from the installation of an efficient clothes washer. Thus, it is useful to look at whether the results correlate with the additional measures installed at each site. To do this, we set up a form of a Conditional Demand Model (CDM). In this case, we establish dummy variables to represent the presence of measures and seek to find a useful correlation between the amount of savings and the measures. We considered a regression against the estimated savings or the size of the measures but this approach is not wise for several reasons. If the savings estimates are subject to error and variability, it is known that a regression will produce biased coefficients. Furthermore, the measure size descriptions appeared to often be unreliable. For example, window area appeared to be the only measure size variable that was reliable and there were insufficient cases for a useful analysis of windows.

Our final selection of CDM variables was limited to clothes washer, furnace, or weatherization dummy variables. None of these variables proved to be statistically significant. As shown in Even though the CDM estimate does not show statistically significant impacts for other measures, it does seem more satisfying to recognize that other measures are expected to impact annual savings results. Thus from Table 8, our best estimate of savings is 71.1 ± 23.3 annual therms.

Table 8, the Tankless Water Heater is itself a significant indicator of savings but none of the other measures provide significant results. The resulting CDM estimate is almost identical to the mean savings shown in Table 4 so it becomes a matter of choice whether to use the CDM estimate or not.

Even though the CDM estimate does not show statistically significant impacts for other measures, it does seem more satisfying to recognize that other measures are expected to impact annual savings results. Thus from Table 8, our best estimate of savings is 71.1 ± 23.3 annual therms.

Table 8. Conditional Demand Model Results, Annual therms Saved

Model Type	Variable	Mean	Standard Error	t-Statistic	P Value (2-tailed)	Count	95% CL Lower	95% CL Higher
Water Heating Model	Tankless Water Heater	73.6	11.6	6.4	0.000	197	50.8	96.4
	Clothes washer	-31.0	40.6	-0.8	0.446	197	-111.0	49.1
Space Heating Model	Tankless Water Heater	69.1	24.0	2.9	0.004	209	21.8	116.3
	Clothes washer	-3.1	33.3	-0.1	0.926	209	-68.8	62.6
	Furnace	19.2	49.8	0.4	0.700	209	-79.0	117.4
	Weatheri-zation	15.5	67.8	0.2	0.819	209	-118.2	149.2
Both Models	Tankless Water Heater	71.1	11.9	6.0	0.000	406	47.8	94.5
	Clothes washer	-3.1	23.2	-0.1	0.893	406	-48.8	42.5
	Furnace	21.0	39.7	0.5	0.598	406	-57.1	99.1
	Weatheri-zation	-9.9	38.4	-0.3	0.796	406	-85.5	65.6