Efficiency Vermont's
Home Performance with ENERGY STAR® Program
Impact Evaluation
Final Report

Prepared for the

Vermont Department of

Public Service

June 3, 2013

Prepared by

West Hill Energy and Computing, Inc.

with GDS Associates

Table of Contents

List of Tables

List of Appendices

Executive Summary

This report provides a detailed description of the impact evaluation conducted for Efficiency Vermont's (EVT's) Home Performance with ENERGY STAR® Program (HPwES) for program years 2008 to 2010. It is the first comprehensive impact evaluation of EVT's residential retrofit programs and provides a benchmark for future program and evaluation activities. Billing analysis was used for both electric savings and savings associated with unregulated fossil fuel, mainly #2 home heating oil and propane. A participant survey also informed the results of the unregulated fossil fuel component of the study.

Overall, evaluation results show the HPwES program is saving both fossil fuels and electricity. Results indicate that HPwES participants saved approximately 16.5 MMBtu and 384 kWh annually (excluding savings from cord wood or wood pellets). For fossil fuels, the evaluation results are based on 82 homes primarily consisting of insulation and air sealing measures. Results indicate that participants achieved savings of about 18% of home heating fuel consumption, on average. HPwES quality control procedures were improved after the study time frame and savings from these program enhancements may be quantified in future evaluations.

There is always some uncertainty associated with the results of evaluation studies and this report is no exception. Every effort was made to reduce uncertainty and provide the most reliable evaluation results for both the fossil fuel and electric savings. However, there are many factors that influence energy consumption and not all can be precisely measured. Sources of uncertainty discussed in Section 4.3.1 include, but are not limited to, sample size, secondary heating (including wood stoves), irregularity of fuel deliveries, and non-response.

Key findings indicate EVT's HPwES realization rate[1] of 86% +/- 12% for electric savings was found to be substantially higher than indicated by other recent impact evaluations of similar residential programs in the region. In addition, the realization rate of 51% +/- 13% for fossil fuels was found to be within the range of some of the realization rates for natural gas savings found for other, similar programs. While recognizing that residential programs may differ in terms of delivery methods, housing stock and market characteristics, a comparison of realization rates from other residential programs in the region are discussed in Section 4.5 of this report for informational purposes. This report also provides recommendations for improving the HPwES program as well as future evaluations.

ES-1. Evaluation Approach

The purpose of this impact evaluation is to establish the first year gross energy savings from electricity and unregulated fossil fuels for installations in program years 2008 through 2010. This evaluation was designed to estimate the savings realization rate (RR), i.e., the ratio of the evaluated gross savings to the HPwES program reported gross savings. Electric savings were

[1] The realization rate is the ratio of evaluated savings to program report savings. A realization rate of 75% indicates that the impact evaluation found 75% of the program reported savings were actually achieved.

verified through a fixed effects billing analysis. Verified unregulated fossil fuel savings were estimated based on annualized consumption. All results were weather normalized as appropriate. No sampling was necessary, and thus, the results were not affected by sampling error.

Due to lack of regulation in the bulk fuel delivery process, evaluating savings from these heating fuels poses challenges that go beyond the issues encountered in the verification of savings from regulated energy sources, such as electricity or natural gas. The hurdles in this part of the analysis included obtaining participant's consent and usage data from fuel dealers. Seventy-three percent of participants also had secondary heat sources, mainly wood, that also complicated the analysis.

The fossil fuel component of this evaluation included the following elements:

1. A participant screener survey to obtain permission to request billing records from fuel dealers and collect additional information about heating consumption patterns
2. Requests to fuel dealers to provide the billing records
3. Compilation of the survey results, program and billing records to conduct the billing analysis

ES-2. Vermont Home Performance with ENERGY STAR Program Description

The primary goal of the Vermont Home Performance with ENERGY STAR Program (HPwES) is to encourage homeowners to use certified home performance contractors to identify and implement improvements that will reduce energy costs and use. Efficiency Vermont offers up to $2,000 in incentives per household to help Vermonters pay for energy efficiency home improvements completed by a certified Home Performance with ENERGY STAR contractor. Efficiency Vermont also provides customer service, contractor training and technical assistance, and quality assurance.

ES-3. Methods

The electric billing analysis was conducted using a cross sectional, time series (CSTS) regression, interrupted at the time of the measure installations. Weather effects and EVT measure installations were included as predictor (independent) variables and the response (dependent) variable was the daily energy consumption. The regression coefficients for program variables were used to estimate the program savings.

A component of the modeling process was to compare alternative models to determine the model that best fits the data and to assess the relative importance of specific variables or groups of variables. Standard statistics, such as the coefficient of determination (R2) and T-values for specific parameters, were compared. Diagnostics were utilized to assess the validity of the regressions model. The regression analysis was conducted at the household level and by major measure group. To estimate the total program evaluated savings, the realization rates from the billing model were applied by measure group.

The fossil fuel billing analysis used a normalized annual consumption model. Given the few number of homes in the model (82), savings could be estimated at the houshold level only. Due to the complexities of the unregulated fossil fuel market, a screener survey was fielded to obtain permission to request billing records from the fuel dealers. This survey also included questions about household characteristics that would provide necessary insight for the fossil fuel billing model such as questions pertaining to current and past use of secondary heat, changes in occupancy, schedules and vacations and the presence and use of heating equipment and other energy-intensive appliances. To boost the response rate and miminize non-response bias, an incentive was offered for submitting the signed permission form. The response rate to the survey was unusually high at 52%, resulting in 242 completed surveys.

Attrition occurs in billing analysis for a variety of reasons, including insufficient consumption history. Both the electric and the fossil fuel models have attrition and the number of participants and reasons for their not being included in the models are provided in Tables E-3-1 and E-3-2. In both cases, the reasons for the attrition (such as insufficient billing records) do not relate to the estimated savings and would not be expected to introduce bias into the results.

Table E-3-1: Summary of Attrition in the Electric Billing Model

	Number of Participants	Percent of Participants
Total Participants with Savings from 2008 through 3/31/2011a	1,124	
Participants with No or Insufficient Bills	189	20%
Participants with No Bills	67	4%
Participanst with Other Billing Issues (gaps, negative reads, etc.)	89	9%
Total Participants Removed	345	31%
Total Participants in Billing Analysis	779	69%

a While the evaluation was designed to estimate realization rates for program participants in 2008 through 2010, the sample frame was expanded through the first quarter of 2011 to increase the sample size.

Table E-3-2: Summary of Attrition in the Fossil Fuel Billing Model

Description	Number of Participants	Percent of Participants with Returned Consent and Completed Survey
Participants who returned the consent form and completed the Screener Survey	198	
Participants whose billing records were received	147	74%
Participants with sufficient billing records	82	41%

For the fossil fuel model, participants in the model were compared to all screener survey respondents for some of the key factors that would be expected to affect energy consumption and the participants in the model were found to be similar to all survey respondents. Details of these comparisons are provided in Section 4.

Secondary heat, including wood and pellet stoves, is common in Vermont homes and Efficiency Vermont's program records show that savings are estimated for multiple heating fuels in a home, where appropriate. This evaluation was designed to estimate savings only for electricity and unregulated fossil fuels.

Since EVT recorded the savings for each fuel type on a house-by-house basis, the denominator in the realization rate was carefully calculated to include only the savings associated with the fuels that were also included in the billing model. Thus, the comparison of the model results to the EVT savings was a direct comparison, i.e., EVT calculated the proportion of the savings associated with specific fuels and the model estimated savings for those specific fuels. All fossil fuels were combined in the billing model by Btu, taking into account the Btu content of the various fossil fuels.

It was not possible to estimate savings from the reduction in wood heat, as the uncertainty associated with estimating wood heating use is greater than the magnitude of the expected savings. A detailed analysis was conducted that compared supplemental wood heat use before and after the installation to ensure that the pre/post analysis provided reliable results. This analysis showed that 86% homeowners with wood heat either did not change the way the wood heat was used over the analysis period or used more wood in the post-installation period.[2] Consequently, the presence of wood heat among participants in the model would not be expected to introduce a downward bias to the realization rate.[3]

[2] Participants who reported heating less than 20% the space with wood and/or using the wood heat rarely or only on the few coldest days of the year were omitted from the analysis.

[3] As only fossil fuel billing records were available, homes with increased wood heat use during the post-installation period would have shown lower fossil fuel use, and thus higher savings, than would have been found if a whole house comprehensive analysis could have been conducted.

Three homes included in the billing model had more than one type of fossil fuel for heating, e.g., a central oil boiler and a propane fireplace insert.[4] Billing records for both fuel types were available for one home, and both fuels were used in the analysis. For the other two homes, bills for the primary heating fuel were collected and used in the analysis; billing records for the supplemental heat were not available.

ES-4. Results

The participant survey provided a spectrum of information on heating systems and heating system control strategies, fuel delivery, secondary fuels and demographics. Table E-4-1 provides the percentages of survey respondents that use different types of fuels to heat their homes. As would be expected in the northeast, the predominant fuel is heating oil.

Table E-4-1: Primary Heating Fuel

	All Survey Respondents (n=242)	Percent of All Survey Respondents	Respondents in Billing Model (n=82)	Percent of All Respondents in Billing Model
Oil/Kerosene	143	59%	57	70%
Propane	40	17%	13	16%
Wood/Pellets	52	21%	11	13%
Other[1]	6	2%	0	0%
Refused/Don't Know	1	0%	1	1%

[1] "Other" includes natural gas, electricity and geothermal heat pumps.
a Participants using wood or pellet stoves as a primary heating source who also identified an unregulated fossil fuel as secondary heating source are included in the billing model

Whether a bulk tank is filled when fuel deliveries occur is an important consideration for the fossil fuel billing analysis. As can be seen in Table E-4-2, fuel tanks are generally filled to capacity for the study participants.

[4] As with the wood heat analysis, participants who reported minimal use of supplemental heat were excluded.

Table E-4-2: Delivery Methods

Fuel tank is filled...	All Survey Respondents (n=186)[1]	Percent of Total Respondents	Respondents in Model (n=77)[1]	Percent of Respondents in Model
Most of the Time	166	89%	66	86%
Some of the Time	5	3%	4	5%
Rarely	1	1%	1	1%
Never	4	2%	0	0%
All of the Time	0	0%	0	0%
Refused/Don't Know	10	5%	6	8%

[1] Due to the survey skip pattern, only respondents with an fossil fuel as primary heating fuel during pre or post period are included.

Fossil Fuel Savings

With 82 homes in the final model, the fossil fuel savings could only be estimated at the household level. As is also the case with all participants with fossil fuel savings, insulation and air sealing measures account for 95% of the savings for homes in the final model, indicating that the measure mix is very similar for the homes in the final model and for all program participants with fossil fuel savings.

Overall, the model indicates that HPwES participants are savings approximately 16.5 MMBtu annually from unregulated fossil fuels. Based on the fossil fuel heating use in the homes in the model, this level of savings represents about 18% of the total pre-installation heating consumption. The savings are statistically significant at the 99% confidence level.[5]

[5] While the 90% confidence limits are reported in the table, the savings are statistically significant at a much higher confidence level.

Table E-4-3: Comparison of Program Reported and Evaluated Savings to Annual Energy Use

Average Pre-Installation Fossil Use for Heating[1]	Program Reported Fossil Fuel Savings[2]		Evaluated Fossil Fuel Savings	
MMBtu/ Year/ Household	MMBtu/ Year/ Household	% of Pre-Install Heating Use	MMBtu/ Year/ Household	% of Pre-Install Heating Use
91.2	32.2	35%	16.5	18%

[1] The average pre-installation fossil fuel use is the total household use, including space and water heating and other end uses (such as cooking or clothes dryers in homes with propane).
[2] The HPwES program reported savings were taken from EVT's central program tracking database. The program reported savings for each home were adjusted to account for the specific fossil fuel(s) that were included in the model.

As shown in Table E-4-3 above, for the homes in the model, the program reported savings represent 35% of the pre-installation consumption levels. Comparing the evaluated savings to the program reported savings indicates that about 51% +/- 13% of the reported fossil fuel savings are being achieved.[6]

The low realization rate could be related to a number of factors. Almost all of the program savings (over 90%) are due to improving insulation levels and conducting air sealing. Consequently, the realization rate primarily reflects the accuracy of these savings estimates.

This evaluation was based solely on a billing analysis, which provides a good indicator of overall program impacts but it does not offer insights into the reasons for the gap between program reports and evaluated savings. Future impact evaluation of the HPwES could be designed to include a more detailed review of program procedures, possibly combined with an on-site survey, to investigate the reasons for the low realization rate for fossil fuel heating measures. In addition, the issue of the use of wood as a supplemental fuel source is a complicating factor that should be carefully examined.

Using billing records to determine pre-installation consumption is a key component to making accurate estimates of potential savings, as is consistent with the operational based rating approach as defined by the Department of Energy (DOE). In contrast, the primarily asset-based approach used by the HPwES program relies primarily on the characeristics of the home to estimate consumption levels, which does not fully account for the actual performance of the home and behavioral aspect of energy use. HPwES program tools and procedures for modeling home energy use and estimating savings, regardless of asset or operational rating type, may need to be updated to ensure that past billing data from fuel dealers are being used to calibrate savgings estimates.

Electric Savings

[6] As a validity check, a billing analysis was conducted for 69 homes that received measures through EVT's HPwES program in Vermont Gas's service territory. The realization rate for the natural gas savings was very close to the fossil fuel realization rate of 51%.

The billing analysis was conducted first in its simplest form, which estimated the savings by household, and then with measure groups. The results of the analysis are provided in Table E-4-4 below.

Table E-4-4: Savings by Measure Group from the Electric Billing Model

			Regression Results			
Measure Group	**Number of Homes in the Model**	**Program Savings per Home (kWh/Yr)**	**Evaluated Savings per Home (kWh/Yr)**	**Lower 90% Confidence Limit[1]**	**Upper 90% Confidence Limit[1]**	**Realization Rate**
CFL	314	422	310	305	316	0.74
Refrigerator Replacement	8	823	698	238	1,158	0.85
DHW Conservation	86	416	62	(280)a	405	0.15
DHW Replacement[2]	1	2,972	2,097	N/A	N/A	0.71
HS Fuel Switch[2]	1	3,841	4,722	N/A	N/A	1.23
HS Replacement[2]	1	14,400	12,734	N/A	N/A	0.88
Insulation/Air Sealing	610	254	280	187	374	1.10
Savings per Household	779	451	384	328	439	0.85

1 The confidence intervals reflect the variability in the model. No sampling was required, so there is no sampling error associated with these results.
2 These measures had only one home in the model and, thus, there is no variability with these results. These measures account for a very small percentage of the total program savings.
a Negative numbers are in parentheses. A negative lower confidence limit indicates that the measure-level savings are not statistically different from zero.

For the most part, the realization rates by measure group range between 0.71 and 1.23, indicating that the program estimates of savings are within the expected range. In contrast, the realization rate for DHW conservation measures was found to be 0.15, suggesting that the savings for these measures are being systematically overstated.

The overall summary of evaluated and program reported savings are provided in Table E-4-5. Comparing the savings from the electric billing models to the program-reported savings indicates that about 86% of the electric savings estimated by EVT for the HPwES program were realized, i.e., the program has a realization rate of 86% +/- 12%.[7]

[7] The overall program savings were estimated by applying the realization rates by measure group from the billing model. Due to the measure mix, there is a small difference between the household realization rate from the billing model (85%) and the overall program realization rate (86%).

Table E-4-5: Summary of HPwES Program Reported and Evaluated Savings for Program Years 2008 Through 2010

	Electric Savings (kWh/Yr)	Fossil Fuel Savings (MMBtu/Yr)
Evaluated Savings	349,681	11,476
Lower 90% Confidence Limit	300,933	8.614
Upper 90% Confidence Limit	398,429	14,333
HPwES Program Reported Savings[1]	407,671	22,421
Realization Rate	86%	51%a
90% Confidence Interval[2]	86% +/- 12%	51% +/- 13%
Relative Precision at the 90% Confidence Level[3]	14%	25%

[1] The HPwES program reported savings were taken from EVT's central program tracking database.

[2] The confidenc interval reflects variability in the model, not sampling error. Since there was no sampling for either the electric or fossil fuel models, there is no sampling error associated with the analysis.

[3] Relative precision is the error bound (e.g., 12% for the electric savings) divided by the realization rate. The low value of the realization rate for fossil fuels (51%) contributes to the relative precision of 25%.

a The realization rate for the fossil fuel savings is for heating measures only. However, heating-related measures account for about 98% of the the total program reported savings.

For both the electric and the fossil fuel models, the only consumption records available were for participants, so it was not possible to compare the reduction in energy consumption among participants to an equivalent non-participant group or to incorporate non-participant trend lines into the models. There are two potential consequences of this approach:

1. Savings could be biased upward due to participant inside spillover, i.e., participants may have decided to adopt additional efficiency measures on their own due to positive interactions with the HPwES Program.
2. Some of the change in energy consumption among participants could be due to non-program related impacts that have a widespread effect among homeowners, resulting in a bias of unknown direction.

These effects cannot be disentangled from the evaluated program savings as estimated through the billing analyses.

ES-5. Recommendations & Conclusions

Program Recommendations

Program recommendations are based on information that was gathered or developed through the course of the evaluation and may prove valuable to program implementers in adjusting program procedures to improve the accuracy of savings claims or identify program components that are not performing well. The program recommendations are listed below.

- Obtain written permission to request billing records from fuel dealers as a requirement for participating in the program
  This process would provide valuable information for program staff to assess savings claims and streamline future evaluation efforts.

- Improve tracking and estimation of electric savings from weatherization measures
  Identification of whether the savings are from electric space heat or from the auxillary components of a fossil fuel system would help distinguish homes by heating method and future evaluations might be able to better assess where the savings are occurring and when they are understated.

Evaluation Recommendations

This evaluation represents the first time that there has been a comprehensive, rigorous impact evaluation of one of EVT's residential custom programs. It is also the first time that comprehensive efforts were made to collect billing records from unregulated fuel dealers to conduct a billing analysis of energy savings. This process has provided the context for identifying possible improvements for future evaluation efforts.

- Design future evaluations of HPwES to investigate external influences on energy use
  A study of external influences could incorporate participant inside spillover, which occurs when participants were motivated by their participation in the program to install additional measures outside of the program. The inability to account for external influences may also introduce a bias of unknown direction.

- Investigate the reasons for the low realization rate for fossil fuel heating measures
  Future impact evaluation of the HPwES could be designed to include a more detailed review of program procedures, possibly combined with an on-site survey, to investigate the reasons for the low realization rate for fossil fuel heating measures. In addition, the issue of the use of wood as a supplemental fuel source is a complicating factor that should be carefully examined.

- Continue with impact evaluations of other EVT custom residential programs
  This study and the recent study of VGS Residential Low Income program demonstrate the importance of impact evaluation in identifying potential program implementation issues.

- Conduct studies to assess net-to-gross estimates for EVT's custom residential programs
  There has not been a Vermont-specific independent study of net-to-gross factors for the residential custom sector, and this type of study could provide valuable insights into the mechanisms that create market changes and the factors influencing program participants to take efficiency actions.

- Continue with impact evaluation of unregulated fossil fuels for other programs
  This evaluation shows that it is possible to collect billing records from fuel dealers with careful planning and a sufficient lead time. Designing future evaluations to submit data requests

during the slow periods for heating fuel dealers would be likely to improve the process.*Conclusions*

This study is the first rigorous impact evaluation of EVT's residential programs. The electric billing analysis meets the enhanced rigor standard as defined in the California Energy Efficiency Evaluation Protocols. The fossil fuel billing analysis was successfully conducted for the first time and the normalized annual consumption (NAC) model used to estimate savings meets the basic rigor as defined in the California Protocols. In addition to providing feedback to program implementers, realistic reporting of energy savings provides valuable information for regulators and the public sphere regarding progress toward meeting efficiency goals and establishing the actual magnitude of reductions in greenhouse gas and other emissions that contribute to climate change.

The one potential limitation of this study was that the billing analyses are based only on participants, as it was not possible to collect non-participant billing data within the scope of this project. The program savings estimated from participant models could incorporate net effects, such as participant inside spillover, which may result in an upward bias in the estimated program savings. The inability to account for external influences may also introduce a bias of unknown direction. Future evaluations should be designed to address these issues.

Since the period covered in this evaluation, Efficiency Vermont has made a number of improvements to program implementation, including tightening quality control procedures, raising the rigor of the end-of-year review process, and developing a formal user manual for the software used by contractors to estimate savings. These actions may well constitute the first steps toward achieving higher realization rates in future evaluations.

1 Introduction

This report provides a detailed description and results of the impact evaluation conducted for Efficiency Vermont's (EVT's) residential retrofit programs for projects installed in 2008 through 2010. This study is the first comprehensive impact evaluation of EVT's residential retrofit programs and provides a benchmark for future program and evaluation activities.

Homeowners who participate in the Vermont Home Performance with Energy Star (HPwES) Program commonly use deliverable fuels such as oil, propane, kerosene or biofuel, to heat their homes and program energy savings claims include deliverable (unregulated) fossil fuels as well as electric savings. The impact evaluation measured the savings of both electric and deliverable fuel consumption using participant surveys and billing analysis to establish first year energy savings.

Billing analysis is an effective tool for conducting impact evaluations of retrofit programs as savings are estimated by comparing energy consumption before and after program participation. This approach is particularly effective for evaluating energy savings for electricity, which is delivered via a meter and consumption is recorded on a regular (monthly) basis. Furthermore, billing records are readily available as Vermont's electric utilities send the billing data directly to EVT.

Conducting a billing analysis of deliverable fuels poses multiple challenges not faced when evaluating electricity or natural gas. The unregulated nature of the market greatly increases the difficulty associated with the data collection process. Usage data cannot be obtained from a centralized source, resulting in a need for outreach and coordination efforts with many fuel dealers. While most of the consumption occurs during the winter months, these fuels may be delivered to on site storage tanks on a regular or as needed basis. Consequently, it is more challenging with deliverable fuels to correlate usage to specific time periods.

There are five sections to this report. The remainder of the introduction provides a brief description of the main objectives of the impact evaluation, evaluation approach, followed by a discussion of the context for the evaluation. Section 2 contains a description of the HPwES program and a summary of program accomplishments during that timeframe. Section 3 details the methods, including details about the homes used in the billing analysis models, followed by the presentation of evaluation results in Section 4. Conclusions and recommendations are provided in Section 5.

1.1 Evaluation Objectives

The purpose of this impact evaluation is to establish the accuracy of first year gross energy savings for installations in program years 2008 through 2010. This evaluation was designed to estimate the savings realization rate (RR), i.e., the ratio of the evaluated gross savings to the HPwES program reported gross savings.

The primary method for estimating program savings was a billing analysis of PY2008, 2009 and 2010 participants. Electric savings were verified through a fixed effects billing analysis. Verified unregulated fossil fuel savings were estimated using a normalized annualized consumption (NAC) model. Savings from wood were not verified as part of this evaluation. All results were weather normalized as appropriate. No sampling was necessary, and thus, the results were not affected by sampling error.

1.2 Evaluation Approach

The evaluation strategies were adapted to accommodate the different characteristics and delivery mechanisms of electric and delivered (unregulated) fossil fuels. This section covers the approach to estimating the electric savings, followed by a discussion of the strategy for verifying the delivered fuel savings.

1.2.1 Electric

Electricity consumption is recorded through a meter on a regular (monthly) basis. For homes meeting this criteria a pre/post billing analysis of all participants having sufficient billing records was conducted using a fixed effects regression model with customer-specific intercepts. Measure-level savings were estimated from the billing analysis. This method meets the enhanced rigor as defined in the *California Energy Efficiency Evaluation Protocols*.[8]

1.2.2 Unregulated Fossil Fuels

Heating fuels such as oil, propane and kerosene are purchased in bulk through a largely unregulated energy market. Due to lack of regulation and the bulk delivery process, evaluating savings from these heating fuels poses challenges that go beyond the issues encountered in the verification of savings from regulated energy sources, such as electricity or natural gas.

The unregulated fossil fuel component of the evaluation included the following components:

1. A participant screener survey to obtain permission to request billing records from fuel dealers and collect additional information about heating consumption patterns
2. Requests to fuel dealers to provide the billing records
3. Compilation of the survey results, program and billing records necessary to conduct the billing analysis

The first hurdle was obtaining consent forms from participants to request billing records from the fuel dealers. Consent forms were sent out and a screener survey was conducted to facilitate this process. Participants who did not complete the screener survey or failed to return the consent form were removed from the sample.

The fuel dealers identified by the participants were contacted to request usage records. Based on the cooperation of the fuel dealers, the final sample comprised participants for whom a complete consumption history could be compiled from all fuel dealers used during the analysis period.

[8] California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals. Prepared for the California Public Utilities Commission. April, 2006. Table 1, page 26.

1.3 Evaluation Context and Issues

Billing analysis was selected as an evaluation method for the HPwES program due to the characteristics of the program. Billing analysis is appropriate for retrofit programs where the baseline is the existing condition prior to the installation. In addition, the savings need to be large enough to separate the program effects from the month-to-month variability of residential energy consumption. A general rule of thumb is that billing analysis works when the program savings are expected to be 8% to 10% or more of the total consumption.[9] With larger sample sizes, smaller effects may be estimated.

1.3.1 Bias and Sampling Precision

Bias and sampling precision are two critical factors that affect the underlying reliability of evaluation results. For the billing models conducted for the HPwES program, there is no sampling as all participants with sufficient billing history were included in the models. Thus, a primary concern for this evaluation is the possibility of bias.

The primary sources of potential bias are as follows:

- the final list of participants who have sufficient billing history and are included in the model may not be representative of the entire program population due to attrition
- some external (non-program) influences may affect energy use but cannot be directly included in the regression models
- net effects, such as spillover, may be partially captured in models intended to estimate gross savings
- while it is assumed that the program is strictly retrofit, some measures, particularly heating equipment, may in fact be market opportunity as the equipment has reached the end of its useful life

Each of these potential sources of bias is discussed below.

Attrition

The concern regarding attrition is whether the removal of specific groups of homes with similar characteristics may introduce bias into the regression results. The potential impacts of attrition are dependent upon the relationship between the type of homes removed from the model and the program delivery mechanisms, as well as the methods used to conduct the analysis and calculate evaluated program savings.

For both the electric and fossil fuel components, the analysis method was designed to minimize the impacts of attrition. A critical aspect of the evaluation technique is that post-installation consumption was compared to pre-installation use for each home and house-specific differences consistent across the analysis period are addressed in the analysis. In addition, the results from the billing models were translated to the program as a whole via realization rates, *i.e.*, the relationship between the verified and program claimed savings was used as opposed to

[9] TecMarket Works. *2004 California Evaluation Framework*, prepared for the California Public Utilities Commission and the Project Advisory Group, September 2004, page 101.

estimating average savings per home.[10] Thus, in assessing the potential bias associated with attrition from the billing analysis, the key issue is whether there is any expectation that specific groups of homes have different realization rates rather than whether the homes in the model are a good match to the homes in the population.

Some of the critical factors that are unlikely to be affected by attrition in the context of this study and the applied methods are discussed below.

- Weather effects were directly included in the billing analysis, and thus, there is no reason to expect that there is any impact from attrition associated with different climate zones.
- Fixed characteristics of the homes, such as housing stock, appliance holdings and lifestyle, are addressed through the house-specific analysis method.
- The mix of measures may vary among subgroups of the program participants. In the electric model, the realization rates were determined by major measure group, which should address this issue. For fossil fuels, insulation and air sealing account for the vast majority of savings for the program as a whole and for the participants included in the billing model.

There are two known issues that may potentially introduce bias into the analysis results. If specific large contractors are better or worse at estimating programs savings and if all homes completed by one or more of these large contractors are completely removed from the analysis, attrition could present a potential source of bias. The second potential source of bias comes from participants who tend to move often and are effectively removed from the analysis as a group since there is insufficient pre- or post-installation billing records to be able to estimate savings. There is no way to assess the impact of this effect or modify the billing models to address it. Potential bias from these sources on the realization rate could be either upward or downward.

For the electric model, the attrition was low and most homes with the measures were included. Thus, the model would be expected to be representative of the program population. The attrition is higher for the fossil fuel model and a comparison of characteristics of all survey respondents and those in the model was conducted to assess whether bias could be introduced through attrition.

External Factors

External factors may have as much, or more, impact on energy use than efficiency programs. When savings are estimated from a billing analysis, these external impacts may introduce either an upward or downward bias to the results. Given that the national economy was moving into a period of contraction during the 2008 program year, one would expect that energy consumption may be reduced across the board, making it likely that program reported savings may be under- or over-estimated, depending upon the timing of the pre- and post-installation periods.

---

[10] An alternative strategy would be to use the billing analysis to establish program savings by measure group or by household, apply these values to the program as a whole to estimate the evaluated program savings and then calculate the realization rate for the whole program on this basis. Under this scenario, it would be important to ensure that the participants in the model were similar to the program as a whole in terms of housing stock, weather conditions, and other factors.

The fixed effects and annualized billing models control for the characteristics of the home that are stable over time and also for seasonal changes in energy use that can be directly incorporated into the model, such as weather and monthly or annual variations. However, it is possible that the estimation of program impacts can be affected by other factors that change over time. These types of changes can be conceptualized in two broad categories:

1. Changes in the overall economy that affect the residential market in a global way, such as volatile gasoline prices, unemployment rates, or an increase in home heating costs.
2. Individual changes that affect specific homes, such as acquiring new household members, taking a longer vacation, or having a change in one's work schedule

The most common approach to investigating the external factors is to incorporate non-participant billing data by developing a trend line of non-participant consumption patterns, assessing overall changes in non-participant energy consumption during the same period or directly incorporating non-participants' billing records in the regression models. However, non-participant billing data was not available for this evaluation. It should also be noted that the use of non-participant billing data could introduce net effects, such as free ridership or spillover, into the models. In the end, a billing analysis that includes both participants and a non-participant comparison group will likely produce savings estimates that are somewhere in between gross and net effects and, thus, difficult to interpret with any degree of accuracy.

Many changes occur over time within homes that are completely outside the influence of the program and yet have an impact on energy use within homes. The approach of including all homes with sufficient billing history in the model is intended to provide a sufficient number of homes in the sample to allow these within home variations to balance out. This conclusion was supported by a recent residential impact evaluation which included detailed surveys of homeowners and a restricted billing analysis. This evaluation concluded that modeling the individual changes within each home did not affect the savings estimates.[11]

Given the time and budgetary constraints, homeowner surveys were only conducted for participants in the fossil fuel model for this evaluation. The purpose of these surveys was to obtain permission to verify that homes are heating with unregulated fossil fuels , request billing records from fuel dealers, identify the way that participants used secondary heating sources before and after program services.

Net Effects

The objective of this impact evaluation was to estimate gross savings for the program, i.e., the reduction in energy use at the home directly due to the measures installed through the program. Attributing savings to the program through estimating free ridership and spillover was not a part of this study.

Both the electric and fossil fuel models include participants only, as obtaining non-participant billing data was not feasible with the time line and budget. However, the participant-only

[11] NYSERDA 2007-2008 EMPOWER NEW YORKSM Program Impact Evaluation Report, Prepared for The New York State Energy Research and Development Authority. Prepared by Megdal & Associates, LLC. February 2012.

models may also include net effects in the form of participant inside spillover. This type of spillover occurs when a program participant learns about efficiency through the program and then elects to install additional measures on their own at the same location. Since this net effect occurs in response to information learned through program participation, it can only occur after the program-related installation and will tend to reduce consumption during the post-installation period.

To the extent that participant inside spillover occurs within the analysis period of the billing analysis, it could result in higher savings than are actually achieved directly from the program-related measures. It is possible that effects from participant inside spillover could be incorporated into the estimated gross program savings.

Replacement at End of Equipment Life
Refrigerator, and water and heating system replacements are the measures which could possibly be market opportunity if the existing equipment was at the end of its useful life. While estimating retrofit savings for market opportunity measures would tend to introduce an upward bias into the evaluated savings, there were few installations of these measures during the period covered in this evaluation. Thus, this issue is unlikely to have a substantial impact on the results.

1.3.2 Unregulated Fossil Fuels

In New England, many homeowners use oil, kerosene, propane, biofuel blends, wood, pellets or some combination of these fuels to heat their homes. Conducting an evaluation of savings from these unregulated fuels creates potential issues not commonly encountered with regulated energy sources. Consumption irregularities and secondary heating sources pose particular challenges for this analysis.

Fossil fuel consumption is highly subject to heating season related fluctuations as fuel is largely used for space heating. Some of the issues are as follows:

- Obtaining billing records requires extensive cooperation from participants and fuel dealers.
- fuel deliveries may not occur on a regular basis and the fuel tank may not be filled to capacity during a given delivery
- consumption is dependent on weather and thermostat settings
- secondary heating fuels may be used and consumption patterns are more likely to vary in homes with secondary fuels

The approach to addressing each of these issues is explored below.

Cooperation from Participants and Fuel Dealers
The Impact Evaluation Team anticipated that the process of collecting billing records would be difficult and lengthy, which turned out to be the case. It was necessary to conduct a participant screener survey to obtain permission to request billing records. This step was followed by the requests to the fuel dealers, and, as would be expected, some fuel dealers were more responsive than others. In some cases, there was not a clear and effective method of identifying the correct employee who would be able to fulfill the request, particularly for larges fuel dealers with multiple offices.

Monetary incentives were offered to both participants and fuel dealers to facilitate this process, and internal process were instituted to track all interactions with both parties. Due to delays in the review and fielding of the surveys, the timing of the requests to fuel dealers was moved into the fall, which unfortunately coincided with both the busy season for the fuel dealers and the presidential election.

Irregular Fuel Fuel Deliveries
To be able to determine the amount of fuel used during a specific period from billing records, it is necessary to know the date of the delivery, the amount of fuel delivered and when the tank was filled to capacity. If the fuel tank is filled to capacity at every delivery, this calculation is straightforward. As the level of granularity in the billing records supports only a normalized annual consumption (NAC) model, the pre-installation consumption and degree days were aggregated and compared to the post-installation consumption and heating degree days. Consequently, the inclusion of one or two partial deliveries within the pre- or post-period should not affect the results.

Since identifying partial deliveries is a key aspect of interpreting the billing records, the Impact Evaluation Team collected information from both participants and fuel dealers. Three levels of review were conducted to address this issue.

1. In the screener survey, participants were asked how often the fuel tank was filled to capacity.
2. Fuel dealers were asked to identify partial deliveries, and some fuel dealers were able to comply with this request.
3. The fuel records were scrutinized and homes with regular deliveries of a specific and even number of gallons (such as 50, 100, or 250) were removed.

Homes with many partial deliveries were removed from the billing analysis.

Weather and Thermostat Settings
Weather effects were directly incorporated into the model through the inclusion of heating degree days calculated for each billing period. Since the billing model compared pre- and post-period consumption for each home, thermostat setting would be a confounding factor in the analysis only if participants made changes between the pre- and post-period. The participant screener survey included specific questions to identify these homes and *post hoc* stratification was conducted to assess whether savings were different or more variable in homes with changes in thermostat settings.

Secondary Heating Fuels
Secondary heat is common in Vermont homes and Efficiency Vermont's program records show that savings are estimated for multiple heating fuels in a home where appropriate. The potential savings from heating wood were not estimated in the model. While theoretically evaluating energy savings through a whole house analysis including all heating sources may be desirable, the difficulties with estimating the change in wood use due to energy efficiency measures is highly problematic since wood is purchased in cords, which have a highly variable Btu content ranging from 20 to 40 MMBtu per cord depending on the type of wood and homeowners are unlikely to be able to estimate wood use to greater granularity than half a cord. Thus, the margin

of error in comparing the wood use from one year to the next is likely to be the same order of magnitude as the estimated savings.

The denominator in the realization rate was carefully calculated to include only the savings associated with the fuels that were also included in the billing model. Thus, the comparison of the model results to the EVT savings was a direct comparison, i.e., EVT calculated the proportion of the savings associated with specific fuels and the model estimated savings for those specific fuels. All fossil fuels were combined in the billing model by Btu, taking into account the Btu content of the various fossil fuels.

Billing analysis is a snapshot and changes in the household energy use patterns over time may affect the longer term savings. The screener survey was used to obtain key details about the use of secondary heat sources and changes in the use of primary and secondary heating fuels between the pre- and post-periods. Homes with fuel switches or other unusual heating patterns were removed from the analysis. In addition, *post hoc* stratification was conducted to assess whether the savings in homes with secondary heat were more variable than homes with an alternative heat source. The results are provided for all homes with completed telephone surveys and for homes in the model in Section 4.1.

2 Program Description

Home Performance with ENERGY STAR is a national program, administered by the U.S. Department of Energy, to help homeowners improve the efficiency and comfort of their homes using a comprehensive, whole-house approach, while helping to protect the environment. As the sponsor for the Vermont statewide Home Performance with ENERGY STAR (HPwES) program, Efficiency Vermont's primary goal is to encourage homeowners to use certified home performance contractors to identify and implement energy improvements.

Efficiency Vermont currently offers up to $2,000 in incentives per household to help Vermonters pay for energy efficiency home improvements completed by a certified Home Performance with ENERGY STAR contractor.[12] Customer incentive levels varied substantially during the 2008-2010 period. Prior to fall 2008, Efficiency Vermont offered only reduced interest rate financing to HPwES customers. In fall 2008, Efficiency Vermont introduced a $500 incentive for all completed HPwES projects. In 2009, Efficiency Vermont began offering incentives up to $2,500 for projects completed by Green Mountain customers, using funding from theGreen Mountain Power Energy Efficiency Fund." In 2010, Efficiency Vermont began using Heating and Process Fuel funding to offer incentives up to $2,500 for all completed HPwES projects.

As described on Efficiency Vermont's website, a typical HPwES project starts with a home comprehensive energy audit conducted by a Building Performance Institute (BPI)-certified HPwES contractor. Upon completion of the audit, homeowners are provided with a report outlining suggested home improvements and the associated energy savings. If the homeowner decides to move forward, the HPwES contractor then installs the recommended improvements and tests out the project according to BPI standards. The most commonly installed measures are insulation and air sealing. However, other improvements may include heating system replacement and distribution, domestic hot water or electric efficiency measures.[13]

Between 2008 and 2010, the number of HPwES contractors certified by the Building Performance Institute more than doubled from a total of 40 to a total of 88. The number of completed projects increased from 322 in 2008 to 588 in 2010. Table 2-1 shows the annual program savings reported by EVT for the HPwES program during each of the three program years assessed through this study (2008 through 2010).

HPwES contractors use Efficiency Vermont's Home Energy Reporting Online (HERO) tool to estimate thermal energy savings. Savings for electric measures such as CFLs and appliances are reported based on TRM characterization.

HERO uses an asset-based approach to estimate savings for envelope measures. Savings are estimated based on the existing condition of the building assembly being retrofitted, the ending condition of the assembly, and long term average annual Heating Degree Days (HDD). The asset-based savings estimates are then adjusted using a hybrid approach whereby HDD are adjusted down 60% from the calculation to account for typical internal/ solar gains, modest

[12] http://www.efficiencyvermont.com/for_my_home/ways-to-save-and-rebates/energy_improvements_for_your_home/home_performance_with_energy_star/rebates.aspx
[13] Efficiency Vermont Annual Report 2010, HOME PERFORMANCE REPORT

occupant set-back and shoulder season energy conservation. In addition, an estimate of annual heating fuel consumption is gathered by contractors through an interview with the homeowner. Estimated annual heating fuel consumption is used by contractors and Efficiency Vermont as a "reality check" for screening and identifying high percentage savings estimates and making subsequent adjustments to inputs[14].

Table 2-1: HPwES Program Savings by Year

Program Year	Participants with Installed Measures	Annual Electric Savings (kWh/ per year)	Annual Unregulated Fossil Fuel Savings (MMBtu/year)[1]	Annual Other Thermal Savings (MMBtu/year)[2]
2008	16	10,354	162	23
2009	203	90,404	3,481	807
2010	736	306,859	18,768	5,878
Total	878a	407,617	22,411	6,708

[1] "Unregulated fossil Fuel Savings" include oil, kerosene and propane savings.
[2] "Other Thermal Savings" include wood and natural gas savings.
a Some participants completed projects in more than one year. There were 878 unique participants during this period.

Table 2-2: HPwES Savings by Measure Group for Program Years 2008 to 2010

Measure Type	All Participants with Installed Measures	Annual Electric Savings (kWh/ per year)	Annual Unregulated Fossil Fuel Savings (MMBtu/year)[1]	Annual Other Thermal Savings (MMBtu/year)[2]
Insulation and Air Sealing	682	179,663	21,293	6,631
Heating System Replacement	18	103	351	64
Domestic Hot Water Conservation	99	38,256	80	0
Domestic Hot Water Replacement	17	5,944	78	9
Programable Thermostat	12	30	42	0
CFLs	357	157,790	0	0
Heating System Fuel Switch	1	3,841	0	-13
Refrigerators	9	7,457	0	0
Other[3]	619	14,533	566	17
Total	878	407,617	22,411	6,708

[1] "Unregulated fossil Fuel Savings" include oil, kerosene and propane savings.
[2] "Other Thermal Savings" include wood and natural gas savings.
[3] Other category includes duct sealing and insulation.

Among participants with fossil fuel measures, as shown in Table 2-3, by far the most commonly installed measure group was insulation and air sealing, with 95% of the fossil fuel savings from this measure group.

[14] The average estimated percentage heating energy savings per project, based on contractors' reporting of the fuel consumption gathered by interview with the homeowner, was 32% between 2010 and 2012.

Table 2-3: HPwES Savings by Measure Group for Homes with Fossil Fuel Savings for Program Years 2008 to 2010

Measure Type	Participants with Installed Measures	Annual Fossil Fuel Savings (MMBtu/year)[1]	% of Total Fossil Fuel Savings
Insulation and Air Sealing	587	21,090	95%
Heating System Replacement	17	351	2%
Domestic Hot Water Conservation	58	54	0%
Domestic Hot Water Replacement	16	78	0%
Programmable Thermostat	8	32	0%
Other	534	558	3%
Total	607	22,165	

Since the time period covered in this evaluation, Efficiency Vermont has made a number of improvements to HPwES program implementation, including tightening quality control procedures, raising the rigor of the end-of-year project review process, and developing a formal user manual for the HERO software used by contractors to estimate savings.

3 Methods

This section describes the methods used to estimate evaluated gross savings for both electric and fossil fuels. The subsections below cover the following topics: the electric regression model, including model selection, regression specifics, calculation of savings from regression estimators, regression diagnostics, as well as the screener survey, the fossil fuel model and attrition in the billing models.

3.1 Electric Billing Analysis

The billing analysis was conducted using a cross sectional, time series (CSTS) regression, interrupted at the time of the measure installations. Weather effects and EVT measure installations were included as predictor (independent) variables and the response (dependent) variable was the daily energy consumption. The regression coefficients for program variables were used to estimate the program savings. The following sections describe the regression model, diagnostic techniques, the specifics of the final regression model, and the method for estimating program savings from the regression results.

3.1.1 Electric Regression Model

The model was a generalized linear model with customer-specific intercept of the form shown in the equation below.

$$C_{it} = \alpha_i + \tau_t + \sum_{j=1}^{p} x_{ijt}\beta_j + \sum_{k=1}^{q} z_{ikt}\gamma_k + \varepsilon_{it} \quad , \qquad (1)$$

where

C_{it} is the monthly consumption for the household *i* in period *t*, expressed in monthly kWh per day,

α_i is the "customer-specific" intercept (or error) for household *i*, accounting for unexplained difference in use between households associated with the number of occupants, appliance holdings and lifestyle,

τ_t is the "time-specific" error for period *t*, reflecting the unexplained difference in use between time periods,

x_{ijt} are the predictor variables reflecting the installation of energy efficiency measure *j* for household *i* in period *t,*

β_j are the slope coefficients that quantify the average influence of modeled efficiency measure *j* on monthly consumption,

p is the total number of energy efficiency measures included in the model,

z_{it} are the predictor variables reflecting non-program related effect *k* (such as weather impacts) for household *i* in period *t*,

γ_k represents the slope coefficients that quantify the average influence of modeled non-program related effect *k* on monthly consumption,

q is the total number of non-program related effects included in the model, and

ε_{it} is the error term that accounts for the difference between the model estimate and actual consumption for household *i* in period *t.*

The model used dummy variables, in which the *x*'s for the installed measure groups are one or zero to indicate the installation and the coefficients reflect the savings for the measure group.

A component of the modeling process was to compare alternative models to determine the model that best fits the data and to assess the relative importance of specific variables or groups of variables. Standard statistics, such as the coefficient of determination (R^2) and T-values for specific parameters, were compared.

The information-theoretic approach to model selection was employed to ensure that the selection of the final model is based on objective statistical standards.[15] The information-theoretic approach is designed to allow a group of candidate models to be compared and ranked by use of Akaike's Information Criterion (AIC). The model with the lowest value of the AIC is the one that best fits the data set, *i.e.*, the model that minimizes the information loss. This approach was used in conjunction with a review of the modeling results to ensure that the "best model" in terms of the statistical properties also allowed for improved estimation of the variables of interest.

3.1.2 Regression Specifics

Given the high degree of variability in residential energy consumption, the billing models with a higher number of homes tend to be more reliable than models with fewer homes. For the purposes of maximizing the number of homes in the model, this analysis includes measures installed from the beginning of 2008 through the first quarter of 2011. Table 3-1 provides definitions for the various measure groups assessed in the model.

[15] In billing analysis, the analyst makes many decisions regarding the statistical characteristics of the model and the specific parameters to be included. Thus, there are typically a number of possible models that could be used to estimate savings. The information-theoretic approach provides an objective framework for selecting the best model among a series of competing candidate models. Please refer to *Model Selection and Multimodel Inference* by Kenneth Burnham and David Anderson, Springer-Verlag, NY, 2002.

Table 3-1: Measure Group Definitions

Measure Groups	Description
Lighting	Installation of screw in CFL lamps
Refrigerator Replacement	Replacement of an older, inefficient refrigerator with an ENERGY STAR model
DHW conservation	Installation of low flow and other devices designed to reduce domestic hot water (DHW) use
DHW Replacement	Replacement of an older, inefficient water heater with a high efficiency water heater
Heating System Fuel Switch	Replacement of electric space heat with a high efficiency, alternative fuel heating system
Heating System Replacement	Replacement of older space heating system with a high efficiency heating system
Insulation and Air Sealing	Attic, wall and basement insulation and air sealing

* Insulation and air sealing measures were most commonly installed as a group and it was not possible to estimate the savings for the specific measures.

The variables included in the final measure-level model are described in Table 3-2 below.

Table 3-2: Variables in the Final Model

Measure Group Estimated	Variable Name	Interaction	Description
None	month1 – month56	none	Accounts for time effects not related to the program or other known factors
CFL Bulbs	cflqty	dpost	Number of CFL bulbs installed, set to the quantity of CFL's installed in the post-period (as identified by the dpost variable), zero in the pre-period
Refrigerator Replacement	Dref	dpost	Dummy variable, set to 1 during the post-period if the home received a refrigerator replacement, zero during the pre-period
DHW Conservation	ddhwconlo, ddhwconhi	dpost	Dummy variables, set to 1 if the home received DHW conservation measures, zero during the pre-period; two variables defined by savings level: low savings (under 500 kWh) and high savings (equal to or more than 500 kWh)
DHW Replacement	dhwrep	dpost	Dummy variable, set to 1 during the post-period if the home received a DHW replacement, zero during the pre-period
Heating System Fuel Switch	dhsfs	dpost, heating degree days	Dummy variable, set to 1 if the home received a heating system fuel switch; variable was interacted with heating degree days during the post period, zero during the pre-period
Heating System Replacement	dhsrepl	dpost, heating degree days	Dummy variable, set to 1 if the home received a heating system replacement; variable was interacted with heating degree days during the post period, zero during the pre-period
Envelope (Insulation and/or Air Sealing)	dinsulvlo, dinsullo, dinsulmed, dinsulhi	dpost, heating degree days	Dummy variables, set to 1 if the home received insulation or air sealing; variables were interacted with heating degree days during the post period, zero during the pre-period; four variables defined by savings levels: very low savings (less than 100 kWh), low (100 to under 200 kWh), medium (200 to under 500 kWh) and high (500 kWh or more)
Heating use for homes with a heating system fuel switch	dhsfs	heating degree days	Heating degree days through entire analysis period to account for heating related use in this subset of homes
Heating use for homes with a heating system replacement	dhsrepl	heating degree days	Heating degree days through entire analysis period to account for heating related use in this subset of homes
Heating use for homes with Envelope (Insulation and/or Air Sealing)	dinsulvlo, dinsullo, dinsulmed, dinsulhi	heating degree days	Heating degree days through entire analysis period to account for heating related use in this subset of homes; four variables defined by savings levels
Heating use for all other homes	nhdd	none	This variable contains the heating degree days (base 60°F) across all homes; the other heating degree days variables will reflect the change for the specified subset of homes in comparison to all homes
Cooling use for all homes	ncdd	none	This variable contains the cooling degree days (base 75°F) across all homes; it accounts for cooling-related changes in use

All of the measure group estimators were statistically significant at the 90% confidence level except for DHW conservation measures.

3.1.3 Regression Diagnostics

Part of the analysis was to assess the validity of the regression approach. To this end, a number of diagnostic tests were conducted. The analysis of outliers and influential data points assesses whether a small number of homes have a large influence on the final results. In some cases, this type of analysis can lead to the identification of homes with highly unusual billing patterns, such as might occur with repeated estimated reads. However, in the absence of clear indication that the homes are highly unusual, influential homes are part of the population of program participants and should be kept in the model.

Autocorrelation and heteroskedasticity are common in cross-sectional, time series (CSTS) models and generally affect the estimation of variability in the model rather than creating bias in the estimators. Collinearity created problems in estimating the parameters of interest and can result in estimators of the wrong sign, for example indicating that specific measures increase energy consumption rather than reduce it. These issues are explored further below.

Outliers and Influential Data Points

Outliers and influential data points can be an issue with regression models, particularly if only a small number of households receive a measure of interest. The DFFITS procedure was used to identify outliers and influential data points. This process involves calculating a predicted value two ways, once with a potential influential observation and once without it. If there is a large difference between the two, the case is considered influential. Typically, observations with a value of DFFITS exceeding 2 are considered to be influential. The DFFITS procedure was modified for the CSTS application by summing the DFFITS values by home and identifying outliers as homes with a combined DFFITS of 2 or more.

Heteroskedasticity

Unequal variances result from the wide fluctuations in energy use from one home to the next due to appliance holdings, occupancy and lifestyle, and are exacerbated by anomalous variations in consumption, either due to estimated reads or other unusual circumstances. The inclusion of the customer-specific intercepts does not completely mitigate the unexplained month-to-month variations. Heteroskedasticity can also be a sign of model misspecification, *i.e.*, that critical variables are omitted from the model or extraneous variables are included.

Heteroskedasticity can be detected through plots of the residuals vs. fits and tested by the Goldfeld-Quandt test or other specifications. The Goldfeld-Quandt (GQ) test is particularly useful for assessing heteroskedasticity in CSTS data sets and was used in this evaluation. However, the GQ test does not always distinguish between heteroskedasticity and model misspecification. To assess whether model misspecification was occurring, the GQ test was run twice for models with a high GC statistic, with and without the outliers as defined in the section above. Since the outliers are homes with higher variation and are more likely to have unequal variances and introduce heteroskedasticity into the model, comparing the GQ statistics with and without the outliers should indicate whether a high GQ statistic is due to heteroskedasticity.

Autocorrelation

Autocorrelation is commonly found in time series data, possibly resulting in biased variances. In this model, autocorrelation stems from the pattern of energy consumption during consecutive periods within each home, i.e., the amount of natural gas used in one month is likely to be similar to consumption during the previous month. While the response variable in the fixed-effect model is the deviation from the expected use, this pattern will still hold to some extent.

While a positively autocorrelated data set should produce unbiased estimators, the variances of the coefficients are likely to be smaller than actually supported by the data. A number of strategies for mitigating first-order autocorrelation have been recommended, but even with these alternative strategies, errors are still likely to be understated in autocorrelated data sets, and care should be used in interpreting the results (Ostrom 1990:36).

The Durbin-Watson test is commonly used to assess the presence of first-order autoregression in least squares regression. A related test, the pooled Durbin-Watson test, has been developed for CSTS data and was applied in this evaluation.

Collinearity

Collinearity tends to be an issue whenever many variables are incorporated into the analysis reflecting measures installed at the same time or when other effects have a high correlation with the measure installations. Collinearity results in higher variances for both response and explanatory variables, and sometimes produces estimators having the opposite sign than would be expected. Two approaches to detecting collinearity were pursued:

(1) identifying individual coefficients that are not statistically significant when the F test for overall model is significant, and

(2) reviewing estimators with opposite signs from what would be expected.

In this evaluation, the primary issue with collinearity was that envelope measures (insulation and air sealing) were most commonly installed together and attempting to estimate the savings for the individual shell measures introduced collinearity into the model. The solution was to bundle the measures as a single measure group.

3.1.4 Calculation of Savings from the Regression Estimators

Savings for the non-temperature-dependent measures were estimated by the direct inclusion of a binary variable set to zero during the pre-installation period and one during the post-installation period. The resulting estimators were in units of kWh per day, and were multiplied by 365 days to calculate energy savings per year.

All measures designed to save space heating energy use were modeled by estimating the heating slope for the post-installation period and the heating slope over the entire analysis period; the post-installation variable reflects the difference in heating slopes, and thus the savings. The resulting estimators were in units of kWh savings per degree day, and were multiplied by the annual heating degree days to calculate energy savings per year.

The heating degree days were calculated using weather data collected by the National Oceanographic and Atmospheric Administration (NOAA) from four weather stations reflecting the range of weather patterns in Vermont (Burlington, Barre, Bennington and Saint Johnsbury). Each county was associated with one of these weather stations and weather data from the station was applied to the billing records for all participants with homes in the county.

The daily weather data was collapsed to calculate the heating and cooling degree days for each billing cycle. The mean temperature for the day was calculated by averaging the minimum and maximum temperatures. Heating degree days were calculated from a base of 60°F and cooling degree days from a base of 75°F.

The realization rate for the homes in the model was determined by comparing the program savings from the billing model to the program reported savings. The normalized heating degree savings from 2007 through 2012 from the NOAA weather data were used to estimate the final measure group and program savings.

To estimate the total program evaluated savings, the realization rates from the billing models were applied by measure group. This approach ensured that the final evaluated savings correctly reflected the actual mix of measures in the overall program.

3.2 Unregulated Fuel Billing Model

A general rule of thumb is that billing analysis is most effective when savings are 8 to 10% or more of the overall consumption. Smaller savings can be estimated when there are many homes in the model. Assuming an average annual heating use of 100 MMBtu, homes with claimed savings of less than 8 MMBtu of a fossil fuel were removed from the sample frame, leaving 607 participants in the final sample frame for the screener survey.

3.2.1 Screener Survey

The process of preparing for the fossil fuel billing analysis began with a screener survey.
The primary objective of the screener survey was to obtain permission to request billing records from the fuel dealers. A secondary purpose was to collect information about household characteristics that would provide necessary insight for the fossil fuels billing model. This included questions pertaining to current and past use of secondary heat, changes in occupancy,

schedules and vacations and the presence and use of heating equipment and other energy-intensive appliances.

Due to the focus on obtaining permission to request billing records, an advance letter, consent form and self-addressed, stamped envelope were sent to each participant on the sample frame prior to fielding the survey with the request to sign and return the consent form. The advance letter and consent form are attached as Appendix A.

The screener survey was found to be an effective strategy for bolstering the response rate by increasing the points of contact with program participants. This approach proved successful for participants who either did not recall receiving the consent form, lost the consent form or for whom the program data did not have the correct address. Of the 36 participants who requested a second consent form, 100% returned a completed form in time to be included in our study.

An unexpectedly high percent of participants returned the signed consent form prior to completing the survey. To be included in the model, participants had to return the signed consent form and complete the survey. Outreach to participants who returned the form but had failed to complete the survey proved much less effective. Only approximately 20% of participants in this category were ultimately reached.

During this outreach process, the West Hill Energy heard from some of the participants who returned the consent form but had trouble with completing the survey. Many participants indicated that due to the high volume of political calls occurring during this time period, calls from unknown numbers were not answered. Other participants indicated that unknown calls were only answered when a message was left on their answering machine, however the policy of the survey house was not to leave messages on answering machines. In addition, calls were typically made in the afternoons and evenings, and a few participants indicated that they were not available during those times of day. West Hill Energy worked with those participants who called to schedule a time to complete the survey.

Ultimately, 242 out of the 607 eligible participants in the fossil fuels sample frame completed the screener survey. The sample disposition can be found below in Table 3-3. The consent form and permission form are included as Appendix A and the screener survey instrument is attached as Appendix B. Additional information regarding the collection of the billing records is provided in Appendix C.

Table 3-3: Sample Disposition

Description		Number of Participants in Sample Frame	Percent of Participants in Sample Frame
Total Number of Participants in Sample Frame		607	
Unusable Sample	No Contact Information	71	12%
	Not working/Unusable number	61	10%
Not Contacted	Respondent never available	0	0%
	Answer Machine	107	18%
Unknown Eligibility	No Answer/Busy	13	2%
Not Eligible	Respondent Not Eligible	13	2%
Refusals/Break-off	Refused	81	13%
	Break-off	19	3%
Completed interview		242	40%
Contact rate[a] = ((242+81+19)/(242+81+19+13+107)) = (342/462)		0.740	
Cooperation rate[b] = 242/(242+81+19) = (242/342)		0.708	
Response rate[c] = 242/(242+19+81+107+(0.96*13)) = (242/461.48)		0.524	

a Contact rate = (Completes+refusals+break-offs)/(Completes+refusals+break-offs+not contacted)

b Cooperation rate = Completes/(Completes+refusals+breakoffs)

c Response rate = Completes/[Completes+refusals+breakoffs+not contacted+ (e*(unknown eligibility))], where e is the eligibility rate.

Definitions of Contact Rate, Cooperation rate and Response rate were established by AAPOR (American Association for Public Opinion Research).

3.2.2 Preparing for Calculating Savings

The lumpy nature of unregulated fossil fuel deliveries makes billing analysis more complicated. Some participants receive fairly regular fuel deliveries and others may have the fuel tank filled only when it is almost empty. If the fuel tank is filled to capacity, it is possible to determine the amount of fuel used between deliveries and the billing analysis can effectively estimate savings.

The first part of this process was to assess whether the fuel tanks are filled to capacity with most, if not all, fuel deliveries. This issue was investigated by reviewing information from the both participant screener survey and the billing records received from the fuel dealers, as described below.

1. Participants were asked in the screener survey how often the fuel tank is filled to capacity and almost 90% reported filling the tank most of the time.
2. Fuel dealers were asked to identify whether the tank was filled at each recorded delivery and some fuel dealers were able to provide this information.
3. The delivery amounts were reviewed and it was quite clear when a specific amount of fuel was added to the tank as opposed to filling the tank; these deliveries were easy to identify as the amount delivered was consistently a multiple of 50 gallons, most often 100, 150 or 250 gallons. These homes also matched participants who reported that they regularly had deliveries of specific amounts of fuel.

This analysis supported the results of the participant screener survey, leading to the conclusion that the vast majority of participants fill the tank to capacity at almost every delivery. Another finding of the screener survey is that HPwES participants tend to be more affluent and highly educated than the average Vermonter, which may relate to the high percent that fill the tank every time. Only participants who filled the tank to capacity for most (and usually all) of the deliveries in the analysis were included in the model.

Given the irregular nature of fossil fuel delivery schedules, up to two years of pre and post usage data were used. The pre and post periods for the fossil billing analysis were defined as follows:

- The inner bounds of the pre/post periods were defined as the delivery dates that were closest to the installation date(s) and still fell within the selected period, e.g., the pre-period ended on the fuel delivery date that was before the installation date and closest to the installation date.
- Consumption reported on the delivery date marking the beginning of each installation period was not included in the total annual consumption, e.g., if deliveries were made in March, December, and April of the pre-period, the total fuel consumption for the period was calculated from the December and April deliveries and the heating degree days were calculated for the entire period from the date of the first delivery in March through the April delivery.

Since the vast majority of participants reported that the fuel tank is filled at every delivery, this approach provides an accurate estimate of fuel consumption during the pre- and post-installation periods. This issue was further investigated as part of the data preparation process, as described in Section 4.1.3.

3.2.3 Calculation of Savings

The first step in the analysis was to calculate the heating degree days for each delivery (for the period from the previous to the current delivery) and the aggregate the quantity of fuel installation and the heating degree days to the pre-installation and post-installation level for each home.[16] For homes with the same fuel for water and space heating, the base (non-heating) consumption was calculated as follows:

1. "Base period" deliveries were identified as those deliveries occurring during periods with low or no heating degree days.[17]
2. The base use gallons of fuel per day was calculated for the base period deliveries.
3. The minimum and average (for homes with more than one base period delivery) base use gallons per day was calculated for each home.
4. The total base use gallons were calculated by multiplying the average base use gallons per day by the days in the pre- or post-period and substracting this number from the total gallons used.
5. The calculations were made separately for the pre- and post-periods in homes with changes in hot water fuel between the pre- and post-periods, such as adding a solar water heater or switching fuels.
6. For homes with no base period deliveries, the average of the base use from the homes with base period deliveries was used.

Most homes with the same water and space heating fuel (49 of 51) had base period deliveries that allowed a house-specific estimate of base consumption. One home was eliminated from the final model because the base use as calculated above was greater than the entire fossil fuel consumption.

Once the total heating consumption (in gallons) had been calculated for the pre- and post-periods separate, the next step was to divide by the heating degree days. The pre- and post-period consumption per heating degree day was then compared, and multiplied by the normalized annual heating degree days for the appropriate weather station to determine gallons savings. These values were converted to MMBtu using the MMBtu/gallon conversion factors obtained from the U.S. Energy Information Adminstration.[18]

This approach was adopted to even out the irregular delivery cycles and minimize error due to the potential for incomplete tank fills within the analysis period, *i.e.*, one incomplete fill within the pre-installation period would be added to the other billing records to obtain a complete assessment of pre-installation use. While the survey results and data review indicate that incomplete fills are uncommon, this approach was taken as a precaution.

In addition, the method was verified using the Princeton Scorekeeping Method (PRISM) by conducting a regression of the heating degree days on fossil fuel consumption for the pre-installation and the post-installation period and comparing the results. Base use was estimated

[16] The quantity of fuel in the first delivery was excluded from the aggregation, but the heating degree days were calculated from the date of the first delivery. This process ensured that the quantity of fuel used and the heating degree days were calculated for the same period.

[17] "Low or no heating degree days" was defined as less than 8.3 HDD per day or 250 HDD per month. Heating consumption would be expected to be minimal at these temperatures.

[18] The conversion factors were 138,690 Btu/gallon for #2 fuel oil, 91,333 Btu/gallon for propane and 135,000 Btu/gallon for kerosene.

for homes with water heating of the same fuel type as the space heating through the addition of an intercept in the regression analysis. If the estimated intercept for these homes was in a reasonable range, *e.g.*, the intercept was positive, the heating use was estimated from the model including the intercept; otherwise, the no-intercept model was used. The regression results were reviewed and homes with unreasonable results (such as a difference in use of 100 MMBtu or greater) were removed, resulting in the elimination of five homes.

The heating degree days were calculated using weather data collected by the National Oceanographic and Atmospheric Administration (NOAA) from four weather stations reflecting the range of weather patterns in Vermont (Burlington, Barre, Bennington and Saint Johnsbury). Each county was associated with one of these weather stations and weather data from the station was applied to the billing records for all participants with homes in the county.

The daily weather data was collapsed to calculate the heating and cooling degree days for each billing cycle. The mean temperature for the day was calculated by averaging the minimum and maximum temperatures. Heating degree days were calculated from a base of 60°F. The final estimated savings were based on heating degree days normalized over the five years of available weather data.

3.2.4 Multiple Heating Fuels

Secondary heat, including wood and pellet stoves, is common in Vermont homes and Efficiency Vermont's program records show that savings are estimated for multiple heating fuels in a home, where appropriate. This evaluation was designed to estimate savings only for electricity and unregulated fossil fuels.

Since EVT recorded the savings for each fuel type on a house-by-house basis, the denominator in the realization rate was carefully calculated to include only the savings associated with the fuels that were also included in the billing model. Thus, the comparison of the model results to the EVT savings was a direct comparison, i.e., EVT calculated the proportion of the savings associated with specific fuels and the model estimated savings for those specific fuels. All fossil fuels were combined in the billing model by Btu, taking into account the Btu content of the various fossil fuels.

It was not possible to estimate savings from the reduction in wood heat, as it would have required substantial additional data collection and the uncertainty associated with estimating wood heating use is greater than the magnitude of the expected savings. A detailed analysis was conducted that compared supplemental wood heat use before and after the installation to ensure that the pre/post analysis provided reliable results. This analysis showed that 86% homeowners with wood heat either did not change the way the wood heat was used over the analysis period or used more wood in the post-installation period. Consequently, the presence of wood heat among participants in the model would not be expected to introduce a downward bias to the realization rate.

Three homes included in the billing model had more than one type of fossil fuel for heating, e.g., a central oil boiler and a propane fireplace insert. Billing records for both fuel types were available for one home, and both fuels were used in the analysis. For the other two homes, bills for the primary heating fuel were collected and used in the analysis; billing records for the supplemental heat were not available.

3.3 Attrition in the Billing Models

To conduct a billing analysis, the preferred approach is to include one year of pre- and one year of post-installation billing records for each participant. While this rule of thumb is not immutable, it is important to ensure that critical periods (such as the deep winter for the fossil fuel model) are included in both the pre- and post-periods. In addition, the models work best when the housing units are similar, and thus master-metered multifamily buildings or mobile home parks cannot be included in the analysis. Other residential evaluations have found that about 40% to 60% of the program participants may be eliminated for these reasons.

3.3.1 Attrition in the Electric Model

All participants with sufficient pre- and post-installation billing records were included in the electric billing model. The billing model requires participants with sufficient billing records throughout the pre- and post-installation periods to be able to estimate savings. Data cleaning was conducted to identify homes that could be included in the billing analyses. The first step was to review the billing data provided by EVT and determine the participants with sufficient billing data to be included in the models. This process is described below.

- The billing data was assessed for each participant to ascertain whether there were sufficient pre- and post-installation records for the model. Each participant was required to have at least six months of billing records that covered the winter period before and after the installation of measures.
- The billing data was checked for a variety of anomalies, including gaps in use, negative or zero reads, unusually low consumption (less than would be expected by a refrigerator and a few lights), and excessively lengthy read periods. These homes were generally removed from the model
- Some participants were found to be on multiple rates at the same time. In some cases, such as CVPS's rate 3, multiple reads were combined. A very limited number of participants were eliminated due to an inability to interpret the rate structure.
- The program and billing data were merged to ensure that the participants in the sample frame had electric measures with associated savings.

The results of this process are summarized in Table 3-4 below. Comparing the number of participants with billing data to the participants determined to be eligible for inclusion in the billing analysis indicates an attrition rate in the range of about 30%. This result is substantially better than other impact evaluations based on billing analysis in the residential sector, including the recent impact evaluation completed by Energy and Resource Solutions (ERS) for NYSERDA on the Con Edison and National Grid natural gas efficiency programs (June, 2010).[19]

[19] Also see "Impact Evaluation of the 2005 California Low Income Energy Efficiency (LIEE) Program," prepared for SCE, PG&E, SDG&E and Southern California Gas by West Hill Energy and Computing, Inc, August, 2008. See Chapter 4.

Table 3-4: Summary of Attrition in the Electric Billing Model

Description	Number of Participants	Percent of Participants
Total Participants with Savings from 2008 through 3/31/2011	1,124a	
Participants with No or Insufficient Pre-Installation Bills	140	15%
Participants with No or Insufficient Post-Installation Bills	49	5%
Participants with No Bills	67	4%
Participanst with Other Billing Issues (gaps, negative reads, etc.)	89	9%
Total Participants Removed	345	31%
Total Participants in Billing Analysis	779	69%

a This table includes all projects through the first quarter of 2011. Table 2-1 covers all projects completed through the end of 2010.

3.3.2 Attrition in the Unregulated Fuel Billing Model

Once received, the billing records were reviewed and participants were removed from the billing model for the following reasons:

- billing records reflected water heating and/or cooking use, rather than heating use[20]
- less than four months of billing records or no billing records in the deep winter months of December through February in either the pre- or post-installation period
- change in heating system fuel between the pre- and post-period
- heating tank was not filled to capacity on a regular basis

While all participants in the billing model were required to meet the minimum of four months and some winter months, 95% of participants had at least a year of both pre- and post-installation data.

Of the 198 participants who returned the consent form and completed the survey, as shown in Table 3-5, 82 (41%) were included in the final model, giving an attrition rate is in a similar range to other evaluations using electric or natural gas utility billing data.

Table 3-5: Summary of Attrition in the Fossil Fuels Billing Model

Description	Number of Participants	Percent of Participants with Returned Consent and Completed Survey
Participants who returned the consent form and completed the Screener Survey	198	
Participants whose billing records were received	147	74%
Participants with sufficient billing records	82	41%

[20] This situation occurred most often for participants who use propane for water heating (and/or cooking) and oil for heating. In some cases, the fuel dealers provided only the propane billing records.

The entire sample frame of 539 participants with contact information and MMBtu savings was used to obtain 82 participants with sufficient billing records. There is no sampling error associated with this analysis as the entire sample frame was exhausted in the process of conducting the screener survey and obtaining the billing records.

4 Results

This section covers the results from the various components of the evaluation, starting with the data collected through the participant screener survey. The following section covers the electric billing analysis. The next section is an overview of the evaluation results, followed by a comparison to impact evaluations for other, similar programs for informational purposes.

4.1 Participant Survey

Household characteristics can have an impact on savings in a variety of ways. A change in schedule, prolong period of vacancy, or changes in the number of occupants in a home all have the potential to increase or decrease the energy used for space and water heating in a given period. In addition, many Vermonters use wood or wood pellet stoves as a secondary heat source.

While the primary purpose of the participant screener survey was to obtain permission to request billing records from the fuel dealers, the survey was also used to collect key information about energy consumption patterns. The survey included questions regarding fossil fuel use, secondary heat use, characteristics of fuel deliveries, heating system management and changes in occupancy. A total of 242 of the 607 participants in the sample frame completed the survey.

The telephone survey was a census, that is the entire sample frame was exhausted in the process of completing the 242 telephone surveys. Thus, there is no sampling error associated with the telephone survey results and the only potential concern is non-response bias. However, the response rate was quite high at 52% (242/462), as shown in Table 3-3.[21] While there is always a potential for non-response bias when a survey is conducted, non-response bias is primarily a concern when the potential reasons for non-response are related to the item of interest (in this case, the household savings). For example, non-response bias could be a problem for on-site surveys as some evaluators have speculated that participants who are proud of their homes and pleased with the energy efficiency improvements are more likely to be willing to sign up for the on-site survey.

The telephone survey was designed to minimize non-response bias by offering an incentive for completing the survey and returning the consent form and by making repeated calls to each home. The primary reasons for the non-response are time constraints, screening of calls and availability. There is no evidence to suggest that non-respondents are different from respondents in terms of any specific characteristics of the home.

The results of these queries are discussed below. Since the response rate was high and non-response bias would not be expected to affect the results of the evaluation, comparing the model homes to the survey respondents is a valid mechanism to check whether the model homes are representative of the population. These comparisons are discussed in Section 4.1.

[21] Response rate = Completes/[Completes+refusals+breakoffs+not contacted+ (e*(unknown eligibility))], where e is the eligibility rate. The definition of the response rate was established by AAPOR (American Association for Public Opinion Research).

4.1.1 Heating Fuel Types

The table below provides a summary of primary heating fuels reported by all survey respondents and also for those respondents who were included in the billing model. As shown in Table 4-1, over half of the survey respondents use oil or kerosene as their primary heating fuel, while less than a quarter use either propane or wood fuel.

Table 4-1: Primary Heating Fuels

Primary Heating Fuel	All Survey Respondents (n=242)	Percent of All Survey Respondents	Respondents in Billing Model (n=82)	Percent of All Respondents in Billing Model[1]
Oil/Kerosene	143	59%	57	70%
Propane	40	17%	13	16%
Wood/Pellets	52	21%	11	13%
Other[2]	6	2%	0	0%
Refused/Don't Know	1	0%	1	1%

[1] Participants using wood or pellet stoves as a primary heating source who also identified an unregulated fossil fuel as secondary heating source are included in the billing model.
[2] "Other" includes natural gas, electricity and geothermal heat pumps.

Seventy-three percent of respondents reported using a secondary heat source. As show in Table 4-2, half of those with a secondary heat source cited wood as their fuel type and most participants using wood as a supplemental heat source were heating with a wood heating appliance, such as a wood or pellet stove.

Table 4-2: Secondary Heating Fuels

Secondary Heating Fuel	All Survey Respondents (n=242)[1]	Percent of Total Respondents	Respondents in Model (n=82)	Percent of Respondents in Model
Oil or kerosene	40	17%	11	13%
Propane	38	16%	10	12%
Wood or pellets	89	37%	34	41%
Natural gas	3	1%	0	0%
Electric	22	9%	6	7%
No secondary jeat	65	27%	26	32%
Refused/Don't know	0	0%	0	0%

[1] Multiple responses were allowed and responses reflect all secondary heat use even if it is infrequent or is only used in small area.

A substantial number of respondents (over 20%) reported using a central heating system, such as a furnace or boiler, as a secondary heat source, as shown in Table 4-3 below.

Table 4-3: Secondary Heating System Type

Secondary Heating System Type	All Survey Respondents (n=242)	Percent of Total Respondents	Respondents in Model (n=82)	Percent of Respondents in Model
Furnace or boiler	50	21%	13	16%
Wood or pellet stove	81	33%	32	39%
Wood fireplace	6	2%	2	2%
Space heater	31	13%	9	11%
Gas or propane fire place	24	10%	5	6%

Overall, the participants including in the billing model are quite similar to all survey respondents in terms of primary and secondary heating fuels. Oil or kerosene is the predominant primary heating fuel in both groups, and there was a higher percentage of wood as a primary heating fuel among all participants than found in the billing model, which would be expected as many of these respondents would be likely to be eliminated from the model due to low or no fossil fuel use.

4.1.2 Wood Heat

As wood heat is common in Vermont homes and the evaluation did not estimate the savings associated with wood heat, additional analysis was conducted to assess whether the homes in the model are representative of the population in terms of the incidence of wood heat and changes in the way that the wood heat was used between the pre- and post-installation periods. About 50% of homes in model and in the population use wood for either primary or secondary heat, and 40% use a substantial amount of wood for heating.

Table 4-4: Homes Using Wood for Space Heating

	Surveyed Respondents (n=242)	Respondents in Billing Model (n=82)	Percentage of Surveyed Respondents	Percentage of Respondents in Billing Model
Homes using wood heat	129	41	53%	50%
Homes with substantial wood use[1]	98	33	40%	40%
Primary wood heat	52	11	21%	13%
Secondary wood heat	46	22	19%	27%

[1] The respondent indicated that wood heat is the primary fuel used for heating or that they heated more than 20% of the home with wood (question SUPP3 on the screener survey) and that they used the wood heat more frequently than just on the coldest days of the year (SUPP2).

Changes in the use of secondary wood heat could affect the results of the billing analysis. For example, if participants modify their heating practices to use less wood heat and more fossil fuels during the post-installation period, the fossil fuel savings may be lower than expected. However, as shown in Table 4-5, most participants (almost 60%) reported no change in their secondary wood heat use and a little over a quarter reported using their wood heat more in the post-installation period. Very few participants reported relying more on secondary wood heat after the installation and no respondents reported changing their primary fuel from wood to another fuel. In addition, the homes in the model are very similar to the telephone respondents on this key indicator. This comparison indicates that the homes in the billing analysis are representative of the population of program participants in this respect and the the billing analysis is unlikely to understate program savings.

Table 4-5: Comparison of Changes in Secondary Wood Heating

Comparison of Wood Heat Use during the Pre- and Post-Installation Periods	Surveyed Respondents (n=242)	Respondents in Billing Model (n=82)	Percentage of Surveyed Respondents	Percentage of Respondents in Billing Model
Used wood heat the same	26	13	57%	59%
Used wood heat more in post-installation period	12	6	26%	27%
Used wood heat less in post-installation period	6	2	13%	9%
Unknown[1]	2	1	4%	5%

[1] The question was inadvertently omitted from two surveys.

4.1.3 Delivery Methods

In order to glean insight into the delivery patterns for this particular population, participants were asked about the frequency of deliveries. The results outlined in Table 4-6, show that more than three quarters of participants whose primary heating fuel was delivered in bulk reported having regularly scheduled deliveries while just 18% receive deliveries on an as needed basis. With 82% of the respondents reporting deliveries at regular intervals, the potential for attrition due to unusable consumption history is reduced.

Table 4-6: Delivery Frequency

Delivery Frequency	All Survey Respondents (n=186)[1]	Percent of Total Respondents	Respondents in Model (n=77) [1]	Percent of Respondents in Model
Once a month	69	37%	25	32%
Once every 2 months	39	21%	17	22%
Once every three months	22	12%	12	16%
Once or twice a year	15	8%	9	12%
When needed	33	18%	10	13%
Automatic delivery	3	2%	1	1%
Refused/Don't know	5	3%	3	4%

[1] Due to the survey skip pattern, only respondents with a fossil fuel as primary heating fuel during pre or post period were asked this question.

These survey participants were also asked how often their tank was filled to capacity when deliveries were made as opposed to having a set amount of fuel delivered, *e.g.*, 100 gallons. As can be seen inTable 4-7, 89% of respondents reported that the fuel tanks were filled to capacity most of the time. This single piece of information increases the reliability of the results as it indicates that the fuel delivered on one date is indeed equal to the fuel that was used since the previous delivery.

Table 4-7: Delivery Methods

Fuel tank is filled...	All Survey Respondents (n=186)[1]	Percent of Total Respondents	Respondents in Model (n=77) [1]	Percent of Respondents in Model
Most of the Time	166	89%	66	86%
Some of the Time	5	3%	4	5%
Rarely	1	1%	1	1%
Never	4	2%	0	0%
All of the Time	0	0%	0	0%
Refused/Don't Know	10	5%	6	8%

[1] Due to the survey skip pattern, only respondents with a fossil fuel as primary heating fuel during pre or post period are included.

This information is primarily useful for preparing for the fossil fuel analysis, and comparing the homes in the model to all respondents does not necessarily provide any insights into similarities or differences between the two groups.

4.1.4 Heating System Controls

Another area of interest was how participants control their heating systems. A direct linear relationship between heating degree days and energy consumption is more likely when the thermostat is set at a consistent temperature and left alone. However, as shown in Table 4-8, many Vermonters setback the thermostat, either manually or using a programmable thermostat.

Table 4-8: Heating System Controls

Heating System Controls	All Survey Respondents (n=242)	Percent of Total Respondents	Respondents in Model (n=82)	Percent of Respondents in Model
Set at one temperature and leave	50	21%	15	18%
Manually adjust as needed	108	45%	39	48%
Set back using programmable thermostat	74	31%	25	30%
Other[1]	7	3%	3	4%
Refused/Don't Know	3	1%	0	0%

[1] "Other" includes combinations of set back strategies and zoned systems.

The survey also included questions about how the control of the heating system changed from the pre-installation to the post-installation period. This analysis suggests that most respondents did not change the method of controlling their heating systems. Of the 49 respondents who reported that they set the thermostat at one temperature during the pre-installation period, only six reported reported having installed a programmable thermostat. Likewise, of the 61 respondents with programmable thermostats in the pre-installation period, three responded that they are now leaving the thermostat at one temperature and five reported that they are now manually adjusting the thermostat (see Table 4-9).

Table 4-9: Comparison of Heating System Controls from Pre- to Post-Installation Period

Pre Installation			Post Installation	
Heating System Control	**Number of Respondents**	**Percent of Total Respondents**	**Heating System Control**	**Number of Respondents**
Set at one temperature	49	20%	Manually adjust	1
			Programmable thermostat	6
			No Change	42
Manually adjust	124	51%	Set at one temperature	5
			Programmable thermostat	15
			No Change	104
Programmable thermostat	61	25%	Set at one temperature	3
			Manually adjust	5
			No Change	53

Changes in thermostat setting between the pre and post period could account for a portion of the change in energy consumption that would not necessarily be attributable to the program. However, as shown in Table 4-10, the majority of participants reported no changes to their thermostat setting.

Table 4-10: Changes in Thermostat Settings from Pre- to Post-Installation Period

Description	All Survey Respondents (n=242)	Percent of Total Respondents	Respondents in Model (n=82)	Percent of Respondents in Model
Set Thermostat Higher during Post Period	19	8%	3	4%
Set Thermostat Lower during Post Period	61	25%	23	28%
About the Same	149	62%	55	67%
Refused/Don't Know	13	5%	1	1%

The comparison of all respondents to respondents included in the billing model suggests that the two groups are quite similar in respect to changes in thermostat settings from pre to post period.

4.1.5 Demographics

In order to assess the demographics of the program population, survey participants were asked questions pertaining to the number of occupants in their home, annual income and highest education attainment. All but five participants who completed this survey live year round in their homes and more than three quarters of the residencies had three of fewer occupants.

The sample population proved to be highly educated with 89 % of surveyed participants having attained at least a college degree. Furthermore, 55% of participant's survey reported a post graduate degree as the highest level of education completed by the homeowner. This percentage seems high when compared with the Vermont census data which reports that only 14% of Vermonters age 25 and older have attained a post graduate degree.[22] The median annual income of survey respondents was between $60,000 and $75,000. In contrast, the median household income for the State of Vermont was about $53,000 in 2011.[23]

4.2 Electric Billing Analysis Results

The billing analysis was conducted first in its simplest form, which estimated the savings by household, and then with measure groups. Only homes with installations completed in 2008 through 2010 were included in the models. The results of the analysis are described below.

[22] 2011 American Community Survey 1-Year Estimates
[23] 2011 American Community Survey 1-Year Estimates

4.2.1 Final Regression Model Results

Table 4-9 presents the measure variables, coefficients and t-values for the final regression model including the non-participant trend variable. Estimators in parentheses are negative, which indicate program savings. The percent of total program reported savings for the measure group is provided as an indicator of the importance of the measure. For example, the t-value for CFLs is 4.0, showing that the results are statistically significant for this major measure accounting for 39% of total program reported savings.

Table 4-11: Electric Regression Model Statistics

Measure Group	Number of Homes in Model	Estimator[1]	t-value[2]	Unit of Estimator	Percent of Total Program Reported Savings for Measure Group[3]
CFLs	314	(0.0703)	-4.0	kWh/Day	39%
Refrigerator Replacement	8	(1.9118)	-1.3	kWh/Day	2%
DHW Conservation: Low Savings	63	0.2321	0.5	kWh/Day	9%
DHW Conservation: High Savings	23	(1.2746)	-1.6		
DHW Replacement	1	(5.7451)	-1.5	kWh/Day	1%
Space Heat Fuel Switch	1	(0.8011)	-4.5	kWh/Heating Degree Day	1%
Space Heat Replacement	1	(1.9634)	-10.5	kWh/Heating Degree Day	0%
Insulation/Air Sealing: Very Low Savings	196	(0.0444)	-2.7	kWh/Heating Degree Day	44%
Insulation/Air Sealing: Low Savings	213	(0.0279)	-1.8		
Insulation/Air Sealing: Medium Savings	166	(0.0215)	-1.3		
Insulation/Air Sealing: High Savings	35	(0.2626)	-8.8		
R-squared[4]		0.80			N/A
Total Homes in Model	779				

[1] The "estimator" is the regression coefficient and reflects the impact of the variable on the change in average daily use. A negative value indicates that there are savings for the measure group. A positive value indicates that there is extra use associated with the measure.

[2] The t-value of a regression coefficient measures whether the value of the coefficient is statistically different from zero. The t-statistic is the regression coefficient over the its standard error. If the absolute value of the t-value of 1.64 or higher, the coefficient is statistically different from zero at the 90% confidence level.

[3] This column does not add to 100% as there are a few measures with small savings that were not included in the model.

[4] The R-squared (R^2) measures the proportion of variability in a regression data set that can be explained by the model.

Some of the characteristics of the model and issues with the modeling process are described below. This discussion covers model statistics, issues with modeling DHW conservation and whole house weatherization measures (insulation and air sealing), incorporating heating and cooling effects and checking models for internal consistency.

Model Statistics. The coefficient of determination (R^2), reflecting the proportion of the change in consumption explained by the model, was quite high at 0.80. Fixed effects models generally have a high R^2 as much of the variation is explained by the house-to-house differences reflected in the customer-specific intercepts. Most of the measure group coefficients are of the correct sign and most of these have t-statistics ranging from 2.5 to over 11.

DHW Conservation Measures. Estimating savings from DHW conservation measures can be problematic. In the original model configuration, there were no savings for DHW conservation measures. An alternative approach involved defining two variables for this measure group, one for the homes with low DHW conservation savings (less than 500 kWh per year) and a second variable for homes with high savings. The estimator for homes with low savings was positive, indicating additional use in these homes. The model showed savings for homes with high savings, and the weighted average of the two estimators indicated small average savings for this measure group. Review of the program data indicates that many participants received pipe insulation, and it appears that the savings for this measure are not being achieved.

Insulation and Air Sealing Measures. The electric savings for insulation and air sealing measures were highly variable and the magnitude of the savings was somewhat dependent on the model configuration. A similar approach to the DHW conservation measures was applied, where the homes were divided into four categories based on the level of the estimated savings. The unusual aspect of the results was that the regression-based savings for homes with the lowest program estimated savings were almost five times higher than the program estimated savings. In contrast, the program reported savings for the homes with the highest savings were slightly higher than the model-based estimates.

A possible reason for this counterintuitive result could be related to the different types of equipment used for space heating. While baseboard electric space heat is using electricity directly to heat a home, a fossil fuel boiler also relies on electricity for igniting the burner and pumping the water through the distribution system. These two types of equipment use electricity in different ways and the relationship between the installation of insulation and the reduction in electric use may be different as well. In EVT's tracking system, there is no clear way to differentiate between a reduction in auxiliary electric use and electric baseboard space heat other than the magnitude of the estimated savings.

It must also be noted that the variability in the insulation and air sealing savings from the model could be related to other, unknown factors. In the absence of alternative methods to incorporate non-participant consumptions trends and/or on-site visits to assess house-specific characteristics, there is no additional information to interpret the model results.

Heating and Cooling Effects. Attempts were made to incorporate house-specific heating and cooling effects in the model. This approach allows a separate heating (or cooling) slope for every home in the model and works well with natural gas models in which almost every home uses natural gas for heating. However, only a few homes have electric space heating or cooling, and there are other factors that affect seasonal energy consumption but are not actually related to weather effects, such as increased electric use during the winter due to longer days and more time spent inside or reduced electric use in the summer due to time spent away from home on vacation. It was not possible to identify homes with electric space heat or mechanical cooling from the program data set provided by EVT.

Careful review of the models suggested that the house-specific approach to heating and cooling use was reflecting non-weather dependent effects, and this approach was set aside. For heating measures, the heating slopes were calculated separately for participants in each measure group to improve the estimators. The final model was run with and without a cooling-related variable and the savings were very similar in both configurations of the model.

Comparison of Models. As a final check on the results, the household savings from the final, measure-level model were compared to the more simplistic model, which estimated savings only at the household level. The results from the two different models were very close, indicating that the measure-level model configuration produced a reasonable and consistent estimate of overall program savings.

4.2.2 Measure Group Savings

Savings for specific measure groups are also provided for context and information purposes (Table 4-12). While this type of modeling provides reliable results at the household level, the measure group savings tend to be more variable and less reliable. Although the magnitude of the measure group savings may be variable, the comparison to program reported savings provides some indication of measure with savings that are more difficult to estimate accurately.

Table 4-12: Savings by Measure Group from the Electric Billing Model

			Regression Results			
Measure Group	**Number of Homes in the Model**	**Program Savings per Home (kWh/Yr)**	**Evaluated Savings per Home (kWh/Yr)**	**Lower 90% Confidence Limit[1]**	**Upper 90% Confidence Limit[1]**	**Realization Rate**
CFL	314	422	310	305	316	0.74
Refrigerator Replacement	8	823	698	238	1,158	0.85
DHW Conservation	86	416	62	(280)a	405	0.15
DHW Replacement[2]	1	2,972	2,097	N/A	N/A	0.71
HS Fuel Switch[2]	1	3,841	4,722	N/A	N/A	1.23
HS Replacement[2]	1	14,400	12,734	N/A	N/A	0.88
Insulation/Air Sealing	610	254	280	187	374	1.10
Savings per Household	779	451	384	328	439	0.85

1 The confidence intervals reflect the variability in the model. No sampling was required, so there is no sampling error associated with these results.
2 These measures had only one home in the model and, thus, there is no variability with these results. These measures account for a very small percentage of the total program savings.
a Negative numbers are in parentheses. A negative lower confidence limit indicates that the measure-level savings are not statistically different from zero.

For the most part, the realization rates by measure group range between 0.71 and 1.23, indicating that the program estimates of savings are within the expected range. In contrast, the realization rate for DHW conservation measures was found to be 0.15, suggesting that the savings for these measures are being systematically overstated.

4.2.3 Results of Regression Diagnostics

Using the DFFITS test for outliers described in Section 3.1.3, there were no outliers in the model. Collinearity would not be expected in the model as there was little overlap among major measures. The electric model was found to exhibit both heteroskedasticity and autocorrelation. The pooled Durbin-Watson statistic was 1.1, which is below the cut off of 2.0 indicating that autocorrelation does not exist. The Goldfeld-Quandt statistic for heteroskedasticity was about 11, which is substantially over the F-statistic of 1 required to confirm that the data set is not heteroskedastic. It is also possible that the results of the Goldfeld-Quandt test could indicate that the model could be misspecified, most likely by missing one or more key parameters.

Autocorrelation is likely to artificially reduce the variance, i.e., give the impression that the estimators are more precise than they actually are. Heteroskedasticity is likely to have the opposite effect in that the variances would appear to be larger than they actually are. There is no method to assess the impacts of a possible misspecification in the model.

4.3 Fossil Fuel Billing Analysis

With 82 homes in the final model, the fossil fuel savings could only be estimated at the household level. As is also the case with all participants with fossil fuel savings, insulation and air sealing measures account for 95% of the savings for homes in the final model, indicating that the measure mix is very similar for the homes in the final model and for all program participants with fossil fuel savings.

Overall, the model indicates that HPwES participants are saving approximately 16.5 MMBtu a year due to program measures (Table 4-13). Based on the homes in the model, this level of savings represents about 18% of the pre-installation heating consumption, which is a substantial accomplishment. The savings are statistically significant at the 99% confidence level.[24]

Table 4-13: Comparison of Program Reported and Evaluated Savings to Annual Energy Use

Average Pre-Installation Fossil Use for Heating[1]	Program Reported Fossil Fuel Savings[2]		Evaluated Fossil Fuel Savings	
MMBtu/ Year/ Household	MMBtu/ Year/ Household	% of Pre-Install Heating Use	MMBtu/ Year/ Household	% of Pre-Install Heating Use
91.2	32.2	35%	16.5	18%

[1] The average pre-installation fossil fuel use for heating is estimated annual household use for space heating.
[2] The HPwES program reported savings were taken from EVT's central program tracking database. The program reported savings for each home were adjusted to account for the specific fossil fuel(s) that were included in the model.

For the homes in the model, the program reported savings represent 35% of the pre-installation heating consumption (Table 4-13), on average. Comparing the evaluated savings to the program reported savings indicates that about 51% +/- 13% of the reported fossil fuel savings are being achieved.

Post hoc stratification was attempted to assess the potential impacts of secondary fuels, heating system control methods and changes in thermostat settings. Due to the small sample size, this process did not always provide conclusive results. For the most part, the difference in savings between the strata were not statistically significant. Key findings from the stratification process are described below.

> *Secondary Space Heat:* Homes in the model were almost evenly divided between those homes with wood, electricity or natural gas as an alternative space heating fuel and those without these secondary heating sources. The savings were somewhat lower for the homes without a secondary heat source, but the results are not statistically significant.
>
> *Heating System Controls:* Savings are almost identical for participants who reported setting the thermostat at one temperature and those who reported manual or thermostatic temperature set backs. While the pre-installation heating consumption was more than 10% higher for homes who do not set back the temperature, this difference was also not statistically significant.

[24] While the 90% confidence limits are reported in the table, the savings are statistically significant at a much higher confidence level.

Change in Thermostat Setting: Savings were dramatically lower in the 26 homes who reported changing the thermostat setting, even though most of these participants reported *lowering* their thermostat during the post-installation period (which would be expected to increase savings). This difference was statistically significant, with savings dropping from 19.6 MMBtu/year in the 55 homes with no change in thermostat setting to 12.0 MMBtu/year for the 26 participants who reported modifying the temperature setting on the thermostat. While the pre-installation heating use was lower in the latter group, the same trend was found when comparing the savings as a percent of heating use (25% as compared to 14% for homes with thermostat changes).

To validate these findings, two additional analyses were conducted: 1) an alternative, PRISM-style NAC analysis and 2) a separate billing analysis using billing data provided by Vermont Gas Systems for homes in the VGS service territory that were served through EVT's HPwES Program.[25] The alternative NAC analysis produced average estimated savings virtually identical to the value determined through the primary analysis method. Based on the 76 homes in the VGS territory with sufficient billing records, the billing analysis resulted in a realization rate of 0.47. Both of these alternative analyses support the findings of the unregulated fossil fuel billing analysis.

Overall, the results demonstrate that the HPwES is achieving meaningful fossil fuel savings, butthe realization rate of 51% suggests that a review of the methods for estimating savings may be in order.

4.3.1 Sources of Uncertainty

There is always some uncertainty associated with the results of evaluation studies and this report is no exception. Although every effort was made to reduce uncertainty and provide the most reliable results for this study, there are many factors that influence energy consumption and not all can be precisely measured. Since billing records are essentially a proxy for energy consumption, the results may include measurement error. Billing analysis reflects only the time period included in the analysis and variations over time are outside of the scope of the evaluation. Non-response bias may be introduced through the use of telephone surveys and the process of attempting to collect sufficient billing records. In addition, this evaluation is also subject to uncertainty surrounding changing human behavior and economic conditions which may impact final results.

The fossil fuel savings were particularly difficult to estimate due to a variety factors. While extensive efforts were made to reduce the uncertainty, it was not possible to estimate the savings to the same level of reliability as achieved for the electric savings or for the evaluation of natural gas savings conducted for other utilities. The major sources of uncertainty in relative order of importance are discussed briefly below.

[25] West Hill Energy recently completed the impact evaluation of Vermont Gas's Home Retrofit Program. For this evaluation, VGS provided billing data for all residential program participants, which included EVT's HPwES program participants. The same models used for the impact evaluation of VGS's program were directly applied to the EVT program participants to validate the results of the unregulated fossil fuel analysis.

- Sample Size: There is a high degree of variability in residential heating and the more homes that can be included in the model, the more reliable the results are likely to be.

- Secondary Heating: Many homeowners in Vermont use secondary heating sources, with wood stoves a common choice. Ideally, evaluators would be able to obtain an accurate estimate of all forms of fuel used to heat the home and conduct a whole house analysis. However, this approach is particularly problematic for wood, as it is not possible to estimate the amount of wood used with sufficient accuracy to calculate the savings. In addition, it was not always possible to obtain all fuel records for participants using more than one type of fossil fuel for heating. Changes in the use of secondary heat could affect the results of the analysis.

- Irregularity of Unregulated Fuel Deliveries: Fossil fuel deliveries are often made on an as-needed basis. In contrast, natural gas customers are billed generally on a monthly basis throughout the year, which provides a much richer data set and improves the evaluators ability to separate heating use from year round consumption (due to water heating for example).

- Non-Response and Attrition: Although the telephone survey attempted to reach all participants with sufficient fossil fuel savings, some participants did not complete the survey and evaluators were not able to obtain sufficient billing records for participants who did complete the survey. As there is no direct connection between the expected magnitude of the savings and the reasons these participants were unable to be included in the analysis, it is unlikely that these factors affected the final results.

These issues and the methods used to mitigate the uncertainties are discussed throughout the report.

4.4 Overview of Program Savings

The overall summary of evaluated and program reported savings are provided in Table 4-14. Comparing the savings from the electric billing models to the program-reported savings indicates that about 86% of the electric savings estimated by EVT for the HPwES program were realized, i.e., the program has a realization rate of 86%. For fossil fuels, the realization rate is 51%. The savings from wood, which account for about 20% of the total MMBtu program reported savings, were not evaluated.

Table 4-14: Summary of HPwES Program Reported and Evaluated Savings for Program Years 2008 Through 2010

	Electric Savings (kWh/Yr)	Fossil Fuel Savings (MMBtu/Yr)

	Electric Savings (kWh/Yr)	Fossil Fuel Savings (MMBtu/Yr)
Evaluated Savings	349,681	11,476
Lower 90% Confidence Limit	300,933	8.614
Upper 90% Confidence Limit	398,429	14,333
HPwES Program Reported Savings[1]	407,671	22,421
Realization Rate	86%	51%a
90% Confidence Interval[2]	86% +/- 12%b	51% +/- 13%
Relative Precision at the 90% Confidence Level[3]	14%	25%

[1] The HPwES program reported savings were taken from EVT's central program tracking database.

[2] The confidenc interval reflects variability in the model, not sampling error. Since there was no sampling for either the electric or fossil fuel models, there is no sampling error associated with the analysis.

[3] Relative precision is the error bound (e.g., 12% for the electric savings) divided by the realization rate. The low value of the realization rate for fossil fuels (51%) contributes to the relative precision of 25%.

a The realization rate for the fossil fuel savings is for heating measures only. However, heating-related measures account for about 98% of the the total program reported savings.

b The overall program savings were estimated by applying the realization rates by measure group from the billing model. Due to the measure mix, there is a small difference between the household realization rate from the billing model (85%) and the overall program realization rate (86%).

Since EVT recorded the savings for each fuel type on a house-by-house basis, the denominator in the realization rate was carefully calculated to include only the savings associated with the fuels that were also included in the billing model. Thus, the comparison of the model results to the EVT savings was a direct comparison, i.e., EVT calculated the proportion of the savings associated with specific fuels and the model estimated savings for those specific fuels. All fossil fuels were combined in the billing model by Btu, taking into account the Btu content of the various fossil fuels.

For both the electric and the fossil fuel models, the only consumption records available were for participants, so it was not possible to compare the reduction in energy consumption among participants to an equivalent non-participant group or to incorporate non-participant trend lines into the models. There are two potential consequences of this approach:

1. Savings could be biased upward due to participant inside spillover, i.e., participants may have decided to adopt additional efficiency measures on their own due to positive interactions with the HPwES Program.
2. Some of the change in energy consumption among participants could be due to non-program related impacts that have a widespread effect among homeowners, resulting in a bias of unknown direction.

These effects cannot be disentangled from the evaluated program savings as estimated through the billing analyses.

4.5 Comparison to Other Programs

For comparison purposes, research was conducted into evaluations of similar natural gas programs as described below. The evaluated programs discussed below are all residential, although they differ in terms of the delivery methods and characteristics of the market and all of these studies incorporate some degree of uncertainty. These comparisons are provided for informational purposes to interpret the results of this impact evaluation in the context of the performance of other thermal efficiency programs.

- The MassSAVE program is delivered directly by the utilities and covers many of the same measures included in EVTs HPwES program; the impact evaluation for program year 2006 found that the realization rates of 0.76 for natural gas and 0.47 for electricity.[26]
- Wisconsin's targeted Home Performance with ENERGY STAR program includes the same range of measures as EVT's HPwES program; this program is delivered through home performance contractors. Wisconsin's program was evaluated for program year 2004 and the evaluation was updated in 2009, although the results were still primarily based on the billing analysis conducted for program year 2004. The impact evaluation results indicate the realization rates were 0.44 and 0.54 for natural gas and electricity, respectively, when compared to the savings reported for program year 2004.[27]
- NYSERDA conducted an impact evaluation of its natural gas programs that was completed in June of 2010. This evaluation shows realization rates for the combined residential programs in the range of 0.57 to 0.76 for natural gas and 0.35 to 0.54 for electricity.[28]

In addition, an impact evaluation of the Vermont Gas Residential Retrofit Program was recently conducted for the Vermont Public Service Department.[29] As the VGS and HPwES programs are different in terms of the market served, the type of fuel saved, the program design and the methods for estimating savings[30], and the evaluation methods were modified to address the requirements of the specific programs,[31] the VGS and HPwES realization rates are not directly comparable. This program differs from EVT's HPwES in that it is directly implemented by VGS auditors and pre-installation billing is reviewed to calibrate savings to actual consumption. In addition, VGS targets high use households which are more likely to have cost-effective opportunities for efficiency upgrades. In addition, the VGS program includes a combination of envelope measures (insulation and air sealing) and heating system replacements, whereas the savings from EVT's program are predominantly from envelope measures. The realization rate for natural gas measures installed through the VGS program was found to be 0.89 overall and 0.73 for envelope measures.

---

[26] See references, MASS, Tables 6 and 7 on pdf pages 19 and 20 (report pages 11 and 12).
[27] See references, WIS.
[28] See references, NYSERDA, Table 5-1.
[29] VGS Residential Program Impact Evaluation, Final Report, Prepared for the Vermont Public Service Department. Prepared by GDS Associates with West Hill Energy and Computing, Inc. February, 2013, page 3, Table ES-1.
[30] The VGS program uses an "operational based" rating and the EVT uses an "asset based" rating.
[31] Since the thermal component of the HPwES program saves unregulated fuels such as propane, oil and kerosene, it was more difficult to obtain billing records. In addition, secondary wood is a confounding factor in the estimation of program savings. Overall, the uncertainty is higher for verifying savings from unregulated fuels than natural gas.

EVT's realization rate of 0.86 for electric savings is substantially higher than found in these other impact evaluations. The realization rate of 0.51 for fossil fuels is in the range of some of the realization rates for natural gas savings found for these other, similar programs.

5 Conclusions and Recommendations

This section covers the program recommendations, evaluation recommendations and conclusions.

5.1 Program Recommendations

Program recommendations are based on information that was gathered or developed through the course of the evaluation and may prove valuable to program implementers in adjusting program procedures to improve the accuracy of savings claims or identify program components that are not performing well.

5.1.1 Obtain permission to request billing records from fuel dealers

Applications for the HPwES incentives should include a signed waiver from participants to request billing records from fuel dealers. As EVT expands its incentives for fossil fuel efficiency, making this adjustment to program implementation would facilitate future evaluation efforts.

5.1.2 Review methods of estimating fossil fuel savings for heating measures

The realization rate for the fossil fuel heating measures was found to be 0.51. This result suggests that a review of the methods and inputs for estimating program reported savings for insulation and air sealing measures should be undertaken. Reviewing pre-installation billing records and calibrating savings to pre-installation consumption would be likely to improve savings estimates.

5.1.3 Improve tracking and estimation of electric savings from weatherization measures

The regression results suggest that program reported electric savings, on average, may be understated for these measures. From a billing perspective, there is a difference between installing insulation in a home with a boiler, where the electric savings are from the auxiliary space heat, and upgrading insulation in a home with electric baseboard space heat as primary heat source. If these applications could be assigned different measure codes, it would provide critical information for future evaluations. It would also be helpful to be able to distinguish among homes with boilers, homes with furnaces, homes with heat pump and homes with electric baseboard. With this additional information, future evaluation efforts may be better able to assess where the savings are occurring and when they are being understated.

5.2 Evaluation Recommendations

This evaluation represents the first time that there has been a comprehensive, rigorous impact evaluation of one of EVT's residential custom programs. It is also the first time that efforts were made to collect billing records from unregulated fuel dealers to conduct a billing analysis of energy savings. This process has provided the context for identifying possible improvements for future evaluation efforts.

5.2.1 Design future evaluations to investigate external influences on energy use

The one potential limitation of this study was that the billing analysis results are based only on participants, as it was not possible to collect non-participant billing data within the scope of this project. The program savings estimated from participant models could incorporate participant inside spillover, which occurs when participants were motivated by their participation in the program to install additional measures outside of the program. Even with access to non-participant billing records, it is difficult to disentangle external influences from program impacts. The inability to account for external influences may also introduce a bias of unknown direction. However, this additional information would still provide useful context for interpreting billing analysis results.

5.2.2 Continue with impact evaluations of other EVT custom residential programs

EVT's other custom residential initiatives may also benefit from rigorous impact evaluation. The recent impact evaluation of VGS's Residential Low Income program, completed in tandem with this evaluation of EVT's HPwES program, suggests that the realization rate for the low income program delivered in partnership with other agencies was much lower than the residential retrofit program delivered directly by VGS, which could be partially due to snap back among participants who decide to take their efficiency gains in the form of increased comfort or other unknown factors. This finding points to the importance of impact evaluation in identifying potential implementation issues.

5.2.3 Conduct studies to assess net-to-gross estimates for EVT's custom residential programs

Another potential evaluation that is indicated from this effort is a need to establish net-to-gross factors, such as free ridership and spillover, associated with EVT's custom residential programs. The current net-to-gross factors are negotiated using studies from other jurisdictions or professional judgment. There has not been a Vermont-specific independent study of net-to-gross factors for the residential custom sector, and this type of study could provide valuable insights into the mechanisms that create market changes and the factors influencing program participants to take efficiency actions.

5.2.4 Investigate the reasons for the low realization rate for fossil fuel heating measures

This evaluation was based solely on a billing analysis, which provides a good indicator of overall program impacts but it does not offer insights into the reasons for the gap between program reports and evaluated savings. Future impact evaluation of the HPwES could be designed to include a more detailed review of program procedures, possibly combined with an on-site survey, to investigate the reasons for the low realization rate for fossil fuel heating measures. In addition, the issue of the use of wood as a supplemental fuel source is a complicating factor that should be carefully examined.

5.2.5 Continue with impact evaluation of fossil fuels for other programs

As EVT's programs have expanded to incorporate unregulated fuels, the need for impact evaluation in this market is growing. Although billing analysis with deliverable fuels presents many special challenges, this evaluation shows that it is possible to collect billing records from fuel dealers with careful planning and a sufficient lead time. Designing future evaluations to submit data requests during the slow periods for heating fuel dealers would be likely to improve the process. These strategies could be applied to a variety of other efficiency programs.

5.3 Conclusions

In the past, the PSB's Annual Savings Verification of EVT's portfolio has been the only evaluation of EVT's custom residential programs. Annual Savings Verification is completed within a highly compressed time frame and is restricted by both time and budgetary constraints to a paper review of projects, with some review of billing records for very large projects. This study is the first rigorous impact evaluation of EVT's residential programs, meeting the enhanced rigor standard for the electric savings and basic rigor for the fossil fuel savings as defined in the *California Energy Efficiency Evaluation Protocols*.

As the first rigorous impact evaluation of EVT's HPwES program, this study provides the opportunity to assess the effectiveness of the program and to identify potential problem areas that may need further investigation. The gross electric savings are reliable, partly due to the high percentage of the total number of participants in EVT's program during the analysis period that were included in the billing model. Although the sample size was smaller, the fossil fuel savings were validated through two alternative analyses. Sample error is not a concern with the fossil fuel analysis as attempts were made to contact all eligible participants for the telephone survey.

This report has described the special challenges and sources of uncertainty involved in conducting a billing analysis of deliverable fossil fuels. Nevertheless, this evaluation demonstrates that it is possible to collect billing records from unregulated fuel dealers with careful planning and persistence. This pilot effort has major implications for future impact evaluations as EVT expands its program offerings designed to improve the efficiency of heating with unregulated fossil fuels. In addition to providing feedback to program implementers, realistic reporting of energy savings provides valuable information for regulators and the public sphere regarding progress toward meeting efficiency goals and establishing the actual magnitude of reductions in greenhouse gas and other emissions that contribute to climate change.

The one potential limitation of this study was that the billing analyses are based only on participants, as it was not possible to collect non-participant billing data within the scope of this project. The program savings estimated from participant models could incorporate net effects, such as participant inside spillover, which may result in an upward bias in the estimated program savings. The inability to account for external influences may also introduce a bias of unknown direction. Future evaluations should be designed to address these issues.

Since the period covered in this evaluation, Efficiency Vermont has made a number of improvements to program implementation, including tightening quality control procedures, raising the rigor of the end-of-year review process, and developing a formal user manual for the software used by contractors to estimate savings. These actions may well constitute the first steps toward achieving higher realization rates in future evaluations.

6 References

The California Evaluation Framework. TecMarket Works, et. al. Project Number: K2033910. Prepared for the California Public Utilities Commission and the Project Advisory Group. September, 2004

California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals. Prepared for the California Public Utilities Commission. April, 2006.

MassSAVE Final Summary QA/QC and Impact Study Report. Prepared for Bay State Gas, Berkshire Gas, Cape Light Compact, KeySpan Energy Delivery, National Grid, NSTAR Electric and Gas Corporation, Unitil/Fitchburg Gas and Electric, New England Gas Company and Northeast Utilities/Western Massachusetts Electric. Prepared by RLW Analytics, Middletown, CT. April 8, 2008. http://www.env.state.ma.us/dpu/docs/electric/08-46/82908nsteera6s9.pdf (MASS)

NYSERDA 2007-2008 EMPOWER NEW YORK[SM] *Program Impact Evaluation Report*, prepared for the New York State Energy Research and Development Authority. Prepared by Megdal & Associates, LLC. February 2012.

"Recommendations for Targeted Home Performance with Energy Star Energy Impacts" (Memorandum) from PA Consulting Group to Oscar Bloch, Public Service Commission of Wisconsin. June 28, 2009. http://www.focusonenergy.com/files/Document_Management_System/Evaluation/Recommendations_for_Targeted_Home_Performance_with_ENERGY_STAR_Energy_Impacts_-_Revised_Final_v2.pdf (WIS)

Report on Impact Assessments of Two NYSERDA Natural Gas Efficiency Programs: Con Edison Natural Gas Efficiency Program and National Grid Low Income Natural Gas Efficiency Program, prepared for the New York State Energy Research and Development Authority by Energy Resource Solutions, North Andover, Massachusetts. June, 2010 (NYSERDA)

Sampling: Design and Analysis. Lohr, Sharon L. Duxbury Press, 1999

VGS Residential Program Impact Evaluation Final Report, prepared for the Vermont Public Service Department by West Hill Energy and Computing, Inc., Brattleboro, Vermont and GDS Associates, Inc., Manchester, New Hampshire. May, 2013. (VGS)