VGS Residential Program
Impact Evaluation
Final Report

Prepared for the

Vermont

Public Service Department

April 11, 2013

Prepared by

West Hill Energy and Computing, Inc.

with GDS Associates, Inc.

Table of Contents

List of Tables

List of Figures

Executive Summary

This report provides a detailed description of the impact evaluation conducted for Vermont Gas Systems' residential retrofit programs for projects installed in 2008 through 2010. This evaluation is a comprehensive impact evaluation of VGS's residential efficiency programs and provides a benchmark for future program and evaluation activities.

During the period covered in this evaluation, VGS's efficiency program performance was not under direct regulatory oversight and VGS had discretion to establish its own goals and budgets.[1] The purpose of the study is to verify gross first year savings and use those results to develop realization rates to reflect the portion of the program reported savings that were actually achieved. In addition, this evaluation took place within the context of Public Service Board docket number 7676 investigating the appointment of an entity to provide natural gas efficiency services.

Given the nature of VGS's residential efficiency programs and budgetary constraints of the project, a billing analysis serves as an effective tool for conducting this type of evaluation. Billing models are used to estimate program savings by comparing the energy consumption before and after measure installations. External influences are addressed through the inclusion of a non-participant trend line developed from VGS's non-participant billing records.

This evaluation was designed primarily to verify the impacts of the Residential Retrofit Program. The Residential Equipment Replacement Program was also included as this approach allowed further opportunities for internal validation and could be implemented at a very small incremental cost. The Residential New Construction Program was not included as billing analysis is not an appropriate method for an impact evaluation of this type of program.

ES I. Program Description

VGS offers two major programs for residential existing buildings: the Residential Retrofit Program (RRP) and the Residential Equipment Replacement Program (RER). The RRP has three components, a market rate component (RMR) that is directly implemented by VGS, a low income component operated in conjunction with CVOEO and projects completed by Efficiency Vermont through its Home Performance with Energy Star Program in VGS's service territory. The primary goal of the RRP is to drive natural-gas energy savings through efficiency and building envelope improvements. Typical measures include insulation and air sealing, however heating system and domestic hot water efficiency measures are also offered. Through this program, participants are offered technical assistance, financial incentives and project management to encourage the installation of the efficiency measures. In the RMR component, participants receive a cash incentive ranging from 33 to 50% of the installed cost following completion of the work by a VGS-screened contractor.

[1] VGS files an Annual Report per ruling of the Vermont Public Service Board; this report outlines prior year savings and spending targets as well as future fiscal year budgeting.

The RER Program provides rebates to VGS customers who replace an older furnace, boiler or water heater with a new, high efficiency model. VGS has established fixed rebates depending on the efficiency of the new equipment, based on societal cost effectiveness screening. Low interest financing is offered to in lieu of fixed cash rebates to cover the cost of the equipment. In addition, VGS offers rental high efficiency water heaters to its customers.

While billing models work well for retrofit programs, the RER is a market opportunity program. VGS calculates savings from new standard efficiency equipment based upon federal, state and and/or regional minimum efficiencies, not existing equipment. However, the evaluated savings from the pre/post billing model reflect the reduction in use compared to the existing equipment. This disconnect between the approach used by the program staff to estimate savings and the method used by the evaluators will be likely to result in a higher realization rate that would be achieved if both the program and evaluated savings were estimated on the same basis. As the older equipment in place before the replacement is likely to have a lower efficiency than a new, standard model, the pre/post billing models will overstate savings and the realization rate will be overstated when the pre/post savings are compared to federal standards or regional baselines. Thus, the RER results should be considered an initial estimate and effectively provide an upper bound on the realization rate for this program.

ES II. Methods and Evaluation Approach

A pre/post billing analysis of all participants having sufficient billing records was conducted using a fixed effects generalized linear regression model. Weather effects and VGS 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 and extensive regression diagnostics were conducted.

A non-participant trend line was constructed using non-participant billing records covering the entire analysis period. This group was used to establish overall trends lines in consumption during the study period and to account for widespread changes in energy consumption that could have been affecting both participants and non-participants.

Since the billing model requires participants to have sufficient billing records throughout the pre- and post-installation periods in order to estimate savings, a data cleaning process was used to identify participants for inclusion in the models. Sufficient billing records was defined as having at least six months of billing records that covered the winter period before and after the installation of measures. Approximately half of the eligible participants were found to have sufficient billing records for the regression models.

ES III. Results

As an overview of the program impacts, the energy savings reported by the program and the evaluated savings were compared to the pre-installation average annual energy use by program component, as shown in ES Table 1 below. Since the market rate component of the RRP is specifically targeted toward high use customers, the pre-installation average household use is the

highest of the three program components, at 1,255 therms. The program reported savings indicate that the reduction in annual energy use is about 26% for participants on average, and the evaluated savings are quite similar with a 23% reduction.

ES Table 1: Comparison of Program Reported and Evaluated Savings to Annual Energy Use

	Average Pre-Installation Natural Gas Use[a]	Program Reported Savings		Evaluated Savings	
Program Component	**Therms/ Year/ Household**	**Therms/ Year/ Household**	**% of Pre-Instal Use**	**Therms/ Year/ Household**	**% of Pre-Instal Use**
Residential Retrofit: Market Rate	1,255	326	26%	290	23%
Residential Retrofit: Low Income	882	229	26%	143	16%
Residential Equipment Replacement	980	136	14%	127	13%

a This value reflects the average annual energy use during the pre-installation period for all homes included in the billing models.

The table below (ES Table 2) provides a summary of the results of this impact evaluation by program and component. The billing analysis was conducted first in its simplest form, which estimated the savings by household, and then with measure groups. The final results were based on the measure-group models, averaging the results from the participant only models and the models that included the non-participant trend line. Only homes with installations completed in 2008 through 2010 were included in the models.

ES Table 2: Savings by Measure Group from the Natural Gas Billing Model

			Regression Results			
Measure Group	**Number of Homes in the Model**	**Program Savings per Home (Therms/Yr)**	**Evaluated Savings per Home (Therms/Yr)**	**Lower 90% Confidence Limit[1]**	**Upper 90% Confidence Limit[1]**	**Realization Rate**
Residential Retrofit: Market Rate						
Water Heater Replacement	51	75	152	147	156	2.03
Envelope Measures	202	261	191	190	193	0.73
Heating System Replacement	111	216	228	226	230	1.05
RMR Savings per Household	247	326	290	270	310	0.89
Residential Retrofit: Low Income						
Water Heater Replacement	*10*	*67*	*-4*	*-46*	*38*	*-0.05*
Envelope Measures	131	186	125	112	138	0.67
Heating System Replacement	12	132	98	69	128	0.75
Duct Sealing	53	68	26	9	43	0.37
RLI Savings per Household	132	229	143	128	158	0.62
Residential Equipment Replacement Program						
Water Heater Replacement	834	834	77	41	34	48
Heating System Replacement	1,565	1,565	102	112	108	116
RER Savings per Household	1,654	1,654	136	127	122	132

Italics indicate that the savings are likely to be too small to estimate given the number of homes in the model that received the measure.
[1] The confidence intervals reflect the variability in the modeling, not the sampling precision (as no sampling was necessary).
[a] The estimator is negative, indicating extra use. However, given the few homes with this measure, it is more likely to be an indication that the model was not effective at estimating the savings.

The overall realization rates by program and program component are shown in ES Table 3. Comparing the savings from the billing models to the program-reported savings indicates that about 89% of the savings estimated by VGS for the RMR component of the Residential Retrofit program were realized, i.e., the program has a realization rate of 89%. The analysis shows that the realization rates vary by program, with the RER program having the highest realization rate of 93% and the RLI component with the lowest at 62%.

ES Table 3: Summary of Realization Rates by Program and Program Component

	Residential Retrofit: Market Rate Component	Residential Retrofit: Low Income Component	Residential Equipment Replacement
Realization Rate[1]	0.89	0.62	0.93
Lower 90% Confidence Limit[2]	0.83	0.56	0.89
Upper 90% Confidence Limit[2]	0.95	0.69	0.97
Relative Precision at 90% Confidence[2]	0.07	0.07	0.04
Number of Homes in Model	247	133	1,654

[1] The final estimates of realization rates were developed by averaging the results from the participants-only models and the models including the non-participant trend line.

[2] The confidence limits and relative precision are an indication of the variability in the model. No sampling was conducted for this analysis and all homes with sufficient billing records were included in the models. Thus, there is no sampling error associated with these results.

The RER program has the highest realization rate at 0.93. This value represents an upper bound on the realization rate as it represents the retrofit savings rather than the market opportunity savings.

ES IV. Program Recommendations

These recommendations 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.

Recommendation 1: Work with Champlain Valley Office of Economic Opportunity (CVOEO) to improve the realization rate of the RLI

The RLI component had a realization rate of 62%. While low-income customers may be more likely to "take back" savings from efficiency programs in the form of increased comfort than other segments of the population, it is also possible that there are other contributing factors to the low realization rate, such as the method of estimating savings. For example, it may be worthwhile investigating whether energy savings are compared to the pre-installation energy use to verify that the estimated savings are within a reasonable range.

Recommendation 2: Assess the RER water heater rental measure as part of least cost planning

During the period in this evaluation (2008 to 2010), VGS was apparently claiming savings for renting high efficiency water heaters to its customers. It could be argued that this measure is

required as part of VGS's least cost planning obligations.[2] Providing an analysis to the PSD that demonstrates the savings claimed for rental water heaters exceed the level of efficiency that is mandated by least cost planning would remove any question as to the legitimacy of the savings claimed for this measure.

Recommendation 3: Review the measure characterizations for the self-installed Eco-Kits

The program reported savings for Eco-Kits are greater than claimed for similar measures in other programs both in and outside of Vermont. Savings for showerheads and aerators and other components provided in Eco-Kits should be based on the best available information from other evaluation studies. VGS could either conduct a literature search and adopt values that are supported by evaluation studies, or adopt savings estimates that are equivalent to the estimates for electric savings in the Vermont Technical Reference Manual.

ES V. Evaluation Recommendations

Evaluation of the VGS Residential Program is an important step in insuring that all of the efficiency programs in Vermont meet generally recognized standards with respect to claimed savings. Evaluation protocols such as the *California Energy Efficiency Evaluation Protocols* have been established in other jurisdictions. 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.

Recommendation 1: Continue with impact evaluations of other VGS efficiency programs

This evaluation of the residential retrofit programs represents a comprehensive impact evaluation of VGS's residential retrofit programs. The C&I programs account for a substantial part of VGS's program-reported savings and they represent a part of Vermont's investment in efficiency that has not received the same level of scrutiny that is given to programs operated by Efficiency Vermont and the Burlington Electric Department.

Recommendation 2: Conduct studies to assess net-to-gross estimates for VGS's efficiency programs

Another potential evaluation that is indicated from this effort is a need to establish net-to-gross factors and other factors associated with VGS programs. The change in the estimated savings when the non-participant trend line is included in the models suggests that the net effects, including free riders and spillover, may be substantial. In addition, VGS is claiming savings from Eco-Kits that are installed by the participant and it is possible that some measures from the kits may not be installed. A study designed to establish how often and when these kits are

[2] VGS pointed out that the rental water program is administered through the non-regulated service department within Vermont Gas and VGS customers do have the option to select from a list of certified contractors to supply and install their water heaters.

installed and to develop the baseline would help establish the societal value of this initiative. Similarly a greater understanding of the rental hot water heater, with a focus on baseline and free-ridership, would provide insight into the target population and effectiveness of the program.

ES VI. Conclusion

The gross program savings were estimated through a billing analysis of all participants with complete and reliable billing data, and both internal and external validation was conducted to ensure that the results were within a reasonable range. The internal validations the associated restricted billing analysis to inform the billing analysis, a review of alternative models to incorporate external influences into the billing regression and an assessment of the validity of the model. External validation consisted of comparing the results from the model to other similar programs.

Since this is a comprehensive and rigorous impact evaluation of VGS's residential retrofit and market opportunity programs, it provides the opportunity to assess the effectiveness of the programs and to identify potential problem areas that may need further investigation. For the program components directly implemented by VGS, it appears that VGS is making a serious and effective effort to estimate savings for its residential programs.

A key aspect of VGS's success with the market rate component of the RRP may be its decision to implement an operational based approach to the estimation of program savings. The Department of Energy (DOE) defines operational based rating as a rating approach that takes into account the specific characteristics of the home and lifestyle of its occupants (thermostat setting, etc.) as well as historical energy use. VGS is successfully using this strategy to improve its estimated savings.

The realization rate for the low income program delivered in partnership with CVOEO was much lower, which could be partially due to snap back among participants who decide to take their efficiency gains in the form of increased comfort. However, further investigation into other possible reasons for the lower realization rate should be pursued. It is possible that an operational based approach would be an effective tool to improve estimated savings for low income homes.

1 Introduction

This report provides a detailed description of the impact evaluation conducted for Vermont Gas Systems' (VGS's) residential retrofit programs for projects installed in 2008 through 2010. This evaluation is a comprehensive impact evaluation of VGS's residential retrofit and market opportunity efficiency programs and provides a benchmark for future program and evaluation activities.

During the period covered in this evaluation, VGS's efficiency program performance was not under direct regulatory oversight and VGS had discretion to establish its own goals and budgets.[3] The purpose of the study is to verify gross first year savings and use those results to develop realization rates to reflect the portion of the program reported savings that were actually achieved. In addition, this evaluation took place within the context of Public Service Board docket number 7676 investigating the appointment of an entity to provide natural gas efficiency services.

Natural gas, like electricity, is delivered via a meter and consumption is recorded on a regular (monthly) basis. As thermal energy savings are likely to be significant in relation to total consumption, billing analysis is an effective tool for impact evaluation of retrofit programs. Billing models are used to estimate program savings by comparing the energy consumption before and after measure installations. External influences are addressed through the inclusion of a non-participant trend line developed from VGS's non-participant billing records.

This evaluation was designed primarily to verify the impacts of the Residential Retrofit Program. The Residential Equipment Replacement Program was also included as this approach allowed further opportunities for internal validation and could be implemented at a very small incremental cost. The Residential New Construction Program was not included as billing analysis is not an appropriate method for an impact evaluation of this type of program.

While billing models work well for retrofit programs, the RER is a market opportunity program and the savings should be calculated from new, standard efficiency equipment rather than the existing equipment prior to replacement. As the older equipment in place before the replacement is likely to have a lower efficiency than a new, standard model, the pre/post billing models will overstate savings. Thus, the RER results should be considered an initial estimate and effectively provide an upper bound on the realization rate for this program.

A key component of the evaluation design was to establish the level of granularity, i.e., the number of separate programs and program components to be considered. For example, the residential low income component of the Residential Retrofit Program is delivered through a different mechanism than the market rate component, and it is entirely possible that the realization rates could be different. With input from VGS, the evaluators developed individual realization rates for the low income component (RLI) and market rate components (RMR) of the

[3] VGS files an Annual Report per ruling of the Vermont Public Service Board; this report outlines prior year savings and spending targets as well as future fiscal year budgeting.

Residential Retrofit Program (RRP). A single realization rate was determined for the Residential Equipment Replacement Program (RER).

There are five sections to this report. 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 RRP and RER 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 (SRR), *i.e.*, the ratio of the evaluated gross savings to the VGS program reported gross savings. Savings by major measure group were estimated, providing some insight into whether specific measures groups are more or less likely to achieve the expected savings.

The primary method for estimating program savings was a billing analysis of PY 2008, 2009 and 2010 participants. Verified savings were determined from a fixed effects billing analysis and results were weather normalized as appropriate. No sampling was necessary, and thus, the results were not affected by sampling error.

In addition to the RRP, evaluators also verified the savings of the Residential Equipment Replacement Programs. Since all of the billing and program data was provided, only incremental effort was required to provide verified savings for this other program.

1.2 Evaluation Approach

Natural gas, like electricity, is delivered via a meter and consumption is recorded on a regular (monthly) basis. Since thermal energy savings are likely to be significant in relation to total consumption, a billing analysis is an effective approach for this study. 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*.[4]

A non-participant trend line was constructed using non-participant billing records covering the entire analysis period. This group was used to establish overall trend lines in consumption during the study period and to account for widespread changes in energy consumption that could have been affecting both participants and non-participants, such as changes in the economy and its associated impacts on households in Vermont.

[4] 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 for VGS due to the characteristics of the Program. Billing analysis is appropriate for retrofit programs where energy-intensive equipment is removed and replaced with high efficiency alternatives. 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 10% or more of the total consumption.[5] With a larger sample size, smaller effects may be estimated. VGS's RRP and RER meet these criteria.

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

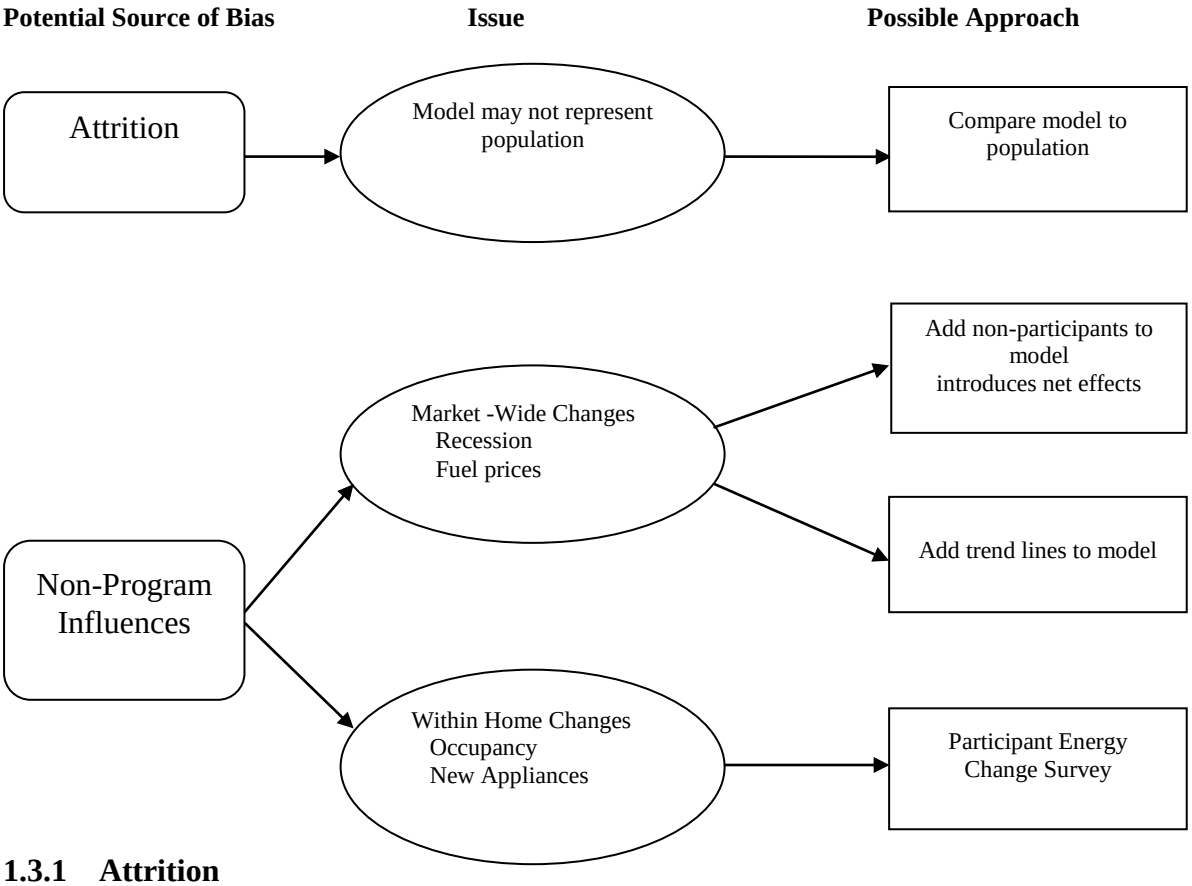
The primary sources of potential bias are as follows:

- 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
- some external (non-program) influences may affect energy use but cannot be directly included in the regression models

The potential sources of bias and strategies for identifying the degree of bias are described in Figure 1-1 below and discussed in more detail in the following paragraphs.

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

Figure 1-1: Sources of Bias and Strategies for Mitigation



1.3.1 Attrition

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 natural gas model) are included in both the pre- and post-periods. The two years of billing history required may not be possible to obtain as some participants may have moved within the analysis period. 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.[6]

The key to ensuring a reliable and defensible billing analysis is the data cleaning process. The purpose of the data cleaning is to identify the homes with complete and reliable billing records to reduce the random error in the model and improve the signal-to-noise relationship. The result of this process is that some homes are removed from the analysis due to insufficient or erratic billing records, which creates attrition. The primary reasons for removal of homes from the billing models are as follows:

- utility billing records are not available

[6] While master-metered units were excluded from the billing analysis, individually metered mobile homes and apartment units with complete billing data were included in the billing models.

- homes do not have sufficient billing records; billing records need to cover critical periods (such as the deep winter for the natural gas model) and include both the pre- and post-measure installation periods
- billing data contain estimated, missing, negative (reconciliation) or highly variable reads; such anomalies create additional error in the model and must be reviewed carefully to assess whether to remove them

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. Two aspects of the selected evaluation methods were designed to minimize the impacts of attrition, as explained below.

- Analysis Method: The estimated program savings were determined for all homes in the regression models. The realization rate was calculated by major measure group by comparing the evaluated savings to the program reported savings for homes in the model with the major measure group. The realization rates were then applied to the program reported savings for all measures to determine the overall evaluation program savings.[7]
- Regression Model: A fixed effects model was used. The fixed effects regression model compares each home to itself, which means that house-specific differences consistent across the analysis period are addressed in the regression analysis.

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 regression models, and thus, there is no reason to expect that there is any impact from attrition associated with the different climate zones.
- Fixed characteristics of the homes, such as housing stock, appliance holdings and lifestyle, are accounted for in the fixed effects model.
- The mix of measures may vary among subgroups of the program participants. Since the realization rates were developed by measure group, any differences in the mix of measures between the participants in the model and those that were removed from the model should not affect the evaluated program savings.

[7] An alternative strategy would be to use the regression results 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 regression analyses were conducted by program and each program has a consistent delivery mechanism, suggesting that one would not expect to see differences in realization rates due to a variety of implementation strategies.

Some participants moved during the analysis period and cannot be included in the analysis since the billing records do not cover the critical months before and after the installation. This type of attrition is random and would not be expected to introduce bias into the results.

1.3.2 External Influences

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 homeowners may reduce energy consumption to save money, making it likely that program reported savings would be under- or over-estimated, depending upon the timing of the pre- and post-installation periods.

The fixed effects model controls 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

Strategies for assessing the impacts of these external factors have been tested in previous billing analyses and will be used in this evaluation, as discussed further in the following sections.

<u>Global External Influences</u>

The primary concern is that relying solely on a billing analysis and failing to consider the broader global energy consumption patterns could result in evaluated program savings that are not necessarily a reflection of program impacts. For example, a billing analysis may show that the savings are substantially lower than estimated by the programs; however, if many homeowners are adding energy consuming equipment and energy consumption is increasing among the general population during the analysis period, this trend may be affecting program participants as well and comparing their pre- and post-installation consumption may not accurately reflect program savings.

There are three common approaches to address the global factors within a statistical billing analysis:

1. include a non-participant comparison group directly in the billing analysis
2. incorporate trend lines based on consumption of the non-participant comparison group

3. incorporate trend lines from third party data on critical market trends, such as the unemployment rate, into the analysis.

Given that the goal of this evaluation was to estimate the realization rates for gross savings, the aim is to develop estimates of evaluated savings that reflect the actual reduction in energy use and do not incorporate net effects such as free riders or spillover. However, eliminating net effects is not straightforward. Incorporating non-participants in the model is a common method of attempting to address global factors, but it also could introduce non-participant spillover.

Non-participant spillover occurs when program participants spread their knowledge of efficiency practices or equipment to non-participants through informal conversations or other mechanisms, or when VGS's general educational efforts are effective at providing the additional knowledge necessary for non-participants to adopt efficiency practices or equipment. VGS operates within a restricted geographical area and has been delivering energy efficiency programs for almost two decades. Consequently, its influence is likely to extend beyond participants.

While the intention of including non-participants in the model is to account for naturally occurring efficiency improvements or trends toward increased use, there is no effective way to separate naturally occurring efficiency from non-participant spillover. Consequently, incorporating non-participants directly into the regression model could add this net effect and reduce the evaluated savings.

However, the participant-only model 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.

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. Using non-participant trend lines should tend to reduce the potential impact on the final results of the billing analysis.

The issue with using external market data is drawing a clear connection between the publicly available trend data and energy consumption. External market data was not pursued in this evaluation due to the time constraints and concerns that the results may not provide insight into the key market effects.

<u>Within Home Influences</u>

Many changes occur over time that are completely outside the influence of the Program and yet have an impact on energy use within homes. Making changes to heating equipment or the addition of a new member to the household are likely to change the patterns of energy use. When conducting a billing analysis, information about these types of changes is not available.

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 the 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.[8] Given the time and budgetary constraints, homeowner surveys were not conducted for this evaluation.

[8] 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.

2 Program Description

VGS offers two major programs for residential existing buildings: the Residential Retrofit (RR) Program and the Residential Equipment Replacement (RER) Program. The RRP is a full service audit program and the RER offers rebates for the installation of high efficiency space and water heating equipment. The next sections provide a brief description of the programs and a summary of the program accomplishments during program years (PY) 2008 through 2010.

2.1 VGS Residential Retrofit Program Description

The RRP has three components, a market rate component (RMR) that is directly implemented by VGS, a low income component operated in conjunction with the Champlain Valley Office of Economic Opportunity (CVOEO) and projects completed by Efficiency Vermont through its Home Performance with Energy Star Program (HPwES) in VGS's service territory.

The primary goal of the RMR component is to address efficiency and building envelope improvements to drive natural-gas energy savings for its residential customers that consume in excess of 50,000 Btus per square foot per year. When appropriate, domestic hot water conservation measures are also installed. Typical measures include insulation and air sealing. Heating system/ heating distribution and domestic hot water efficiency measures are also offered.

Through the RMR component, VGS offers a free energy audit to its customers, which includes a report with recommended efficiency improvements. VGS offers technical assistance, financial incentives and project management to encourage the installation of the efficiency measures. VGS will assign a pre-screened contractor to complete the work at the request of the participant.

Participants are offered a cash incentive of 33% of the installed cost of the measure in homes where the property owner pays the natural gas bills and 50% if tenants are responsible for the bill. Low interest financing is also available to cover the remainder of the cost. VGS implements an operational based approach to the estimation of program savings.[9] This process involves comparing the estimated savings to the actual, pre-installation consumption for each participant.

Low income customers are served through a cooperative agreement with CVOEO). VGS refers interested customers to CVOEO for priority services and contributes a portion of the costs, including income verification and direct measure costs. VGS conducts the cost-effectiveness screening for the measures to be installed in the homes of low income customers. All measures are installed at no cost to the participant.

[9] The Department of Energy (DOE) defines operational based rating as a rating approach that takes into account the fixed characteristics of the home as well as the lifestyle of its occupants (thermostat settings, etc.) and historical energy use. In contrast, the assets-based approach accounts for only for the fixed characteristics of the home and does not address the actual performance.

In addition, VGS works with Burlington Electric Department (BED) and with Efficiency Vermont. VGS assists BED in the implementation of its Time of Sale (TOS) ordinance, which requires rental housing to meet BED's efficiency standard when the property changes hands. VGS offers incentives to BED's TOS participants as is consistent with the non-low income component of the RRP. VGS customers who do not meet the consumption threshold to participate in the RMR component are referred to Efficiency Vermont's Home Performance with Energy Star Program.

VGS reported some changes to program implementation over these three years, as specified below.

- During program year 2008, VGS began providing audit customers with Gas Eco self-installation kits.
- During program year 2009, VGS transitioned to a new software tracking program.
- During program year 2010, VGS improved its coordination with EVT's Home Performance with Energy Star program, allowing VGS customers who participate in EVT's or BED's program to receive VGS incentives for cost effective measures.

These modifications to program implementation would not be expected to have an impact on the realization rate or evaluated savings.

2.2 Residential Retrofit Program Accomplishments

VGS provided the following summary of program accomplishments in its DSM annual reports for program years 2008 through 2010. The savings from the low income component of the program accounted for about 16% of the total program savings in 2010.

Table 2-1: Residential Retrofit Program Reported Savings by Year

Program Year	Audits Completed	Customers with Installed Measures	Program Reported Annual Savings (Mcf per Year)	Completed Low Income Projects *
2010	374	176	8,809	48
2009	436	239	10,348	71
2008	410	164	8,064	53
Total	1,220	579	27,221	172

*Low income projects were completed under a cooperative arrangement between VGS and CVOEO-WX.

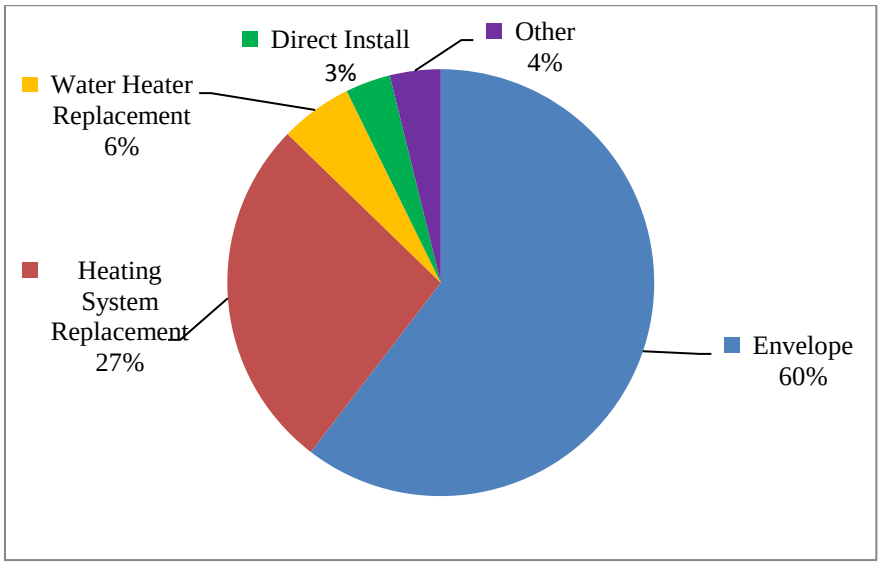
The distribution of savings by measure group is shown in Table 2-2 below and illustrated in Figure 2-1 below. The majority of the annual Mcf savings (60%) are associated with envelope measures such as air sealing and insulation. Heating system replacement is the next highest measure, with 27% of the savings. The remaining 13% of savings is divided among water heater replacements, direct install measures and the other measure.

Table 2-2: RRP Reported Savings by Major Measure Group for PY 2008 to 2010

Major Measure Group	# of Participants with Measure	Program Reported Annual Savings (Mcf per Year)	Program Reported Peak Day Savings (Mcf)	% of Annual Mcf Savings
Envelope	527a	16,446	170.7	60%
Heating System Replacement	196	7,307	77.2	27%
Water Heater Replacement	15	1,492	4.7	5%
Self Install Kit	166	931	3.0	3%
Other[1]	1	1,046	11.5	4%
Totals	579b	27,221	267.1	

[1] The "other" measure was labeled "heat recovery" in VGS's 2010 DSM Annual Report.
a "Envelope" is a combination of the "air infiltration" and "shell" measure categories listed in VGS's DSM annual reports. Many participants received both types of measures. In this table, the number of participants with the measure was assumed to be the higher of the counts in the two measure categories. It may be slightly understated since it is possibly that a few participants installed only shell or only air infiltration measures.
b The total of number participants does not add up the sum of the individual measures as some participants installed more than one measure.

Figure 2-1: Residential Retrofit Program Savings by Major Measure Group



2.3 Residential Equipment Replacement Program Description

Through the RER program, VGS also provides rebates for installing high efficiency space and water heating equipment. VGS describes this program as follows:

> "The Vermont Gas Residential Equipment Replacement Program is designed to encourage customers to purchase and install water and space heating equipment that exceeds both the current standards established by the National Appliance Energy Conservation Act (NAECA), and the de facto baselines in the region. These replacements typically occur when the equipment has failed and can no longer be repaired or has reached the end of its useful life; or when the fuel source for heating a home is being switched to natural gas." (VGS 2010 DSM Annual Report, page RER-1)

In addition, this program includes a water heater rental component, in which "customers and contractors can rent High-efficiency [*sic*] water heaters (.62 energy factor or greater) from VGS." (VGS 2010 DSM Annual Report, page RER-2)

Since the primary method used in this impact evaluation is a billing analysis, the RER program was also evaluated as the additional incremental cost and effort to include the RER program was minimal. While billing models work well for retrofit programs, the RER is a market opportunity program and VGS calculates the savings from new, standard efficiency equipment based upon federal state and/or regional minimum efficiencies, not existing equipment. However, the evaluated savings from the pre/post billing model reflect the reduction in use compared to the existing equipment. This disconnect between the approach used by the program staff to estimate savings and the method used by the evaluators will be likely to result in a higher realization rate than would be achieved if both the program and evaluated savings were estimated on the same basis. Thus, the RER results should be considered an initial estimate and effectively provide an upper bound on the realization rate for this program.

The RER program offers rebates to VGS customers who replace an older furnace or boiler with a new, high efficiency model. VGS has established fixed rebates depending on the efficiency of the new equipment, based on societal cost effectiveness screening. Low interest financing is offered in lieu of fixed cash rebates to cover the cost of the equipment. In addition, VGS offers rental high efficiency water heaters to its customers. The VGS program reported savings are calculated based upon Department of Energy, and local, regional minimum baselines, not existing equipment.

2.4 Residential Equipment Replacement Program Accomplishments

VGS provided the following summary of RER program accomplishments in its DSM annual reports for program years 2008 through 2010 (Table 2-3).

Table 2-3: Residential Equipment Replacement Program Reported Savings by Year

Program Year	Customers with Rebates	Program Reported Annual Savings (Mcf per year)	Program Reported Peak Day Savings (Mcf)
2010	1,932	17,553	139.3
2009	1,600	18,296	175.1
2008	1,517	12,508	123.1
Total	5,049	48,356	437.5

The distribution of savings by measure group is shown in Table 2-4 below. Heating system replacements account for about two-thirds of the program annual savings.

Table 2-4: RER Program Reported Savings by Major Measure Group for PY 2008 to 2010

Major Measure Group	# of Participants with Measure	Program Reported Annual Savings (Mcf per Year)	Program Reported Peak Day Savings (Mcf)	% of Annual Mcf Savings
Heating System Replacement	2,773	31,991	361.1	66%
Water Heater Rental	3,666	16,365	76.4	34%
Totals	5,049a	48,356	437.5	

a The total number of participants does not add up the sum of the individual measures as some participants installed both measures.

3 Methods

This section describes the methods used to estimate evaluated gross savings. The subsections cover the following topics: the regression model, attrition in the billing models, the non-participant trend line, model selection, regression specifics, calculation of savings from regression estimators and regression diagnostics.

3.1 The Regression Model

Weather effects and VGS 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 model was a generalized linear model with customer-specific intercept of the form shown in the equation below.

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

where

C_{it} is the monthly consumption for the household *i* in period *t*, expressed in therms 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. In addition, the information-theoretic approach to model selection was employed to ensure that the selection of the final model is based on objective statistical standards.[10] 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.

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.

3.2 Attrition in the 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 VGS 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 program and billing data were merged to ensure that the participants in the sample frame had natural gas measures with associated savings.

The results of this process are summarized in Table 3-1 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 45% to 50%. This result is similar to 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).[11]

[10] 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.

[11] 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 3-1: Summary of Attrition in the Billing Model

	RMR Program	RLI Program	RER Program	Totals
Total Participants with Savings	467	250	3,220	3,981
Participants with Insufficient Bills	203	101	1,453	1,725
Participants with No Bills	6	5	27	39
Participants with Erratic Consumption Patterns	11	12	86	112
Total Participants Removed	219	118	1,564	1,914
Total Participants in Billing Analysis	247	132	1,654	2,059
% of Participants in Billing Analysis	53%	53%	51%	52%

3.3 Preparation of the Data

The program and billing records were combined to create the data set used for the regression analysis. The program data was collapsed to one record per premise and DSM ID and savings were summed by measure group and a binary variable for each program was set, i.e., if a customer participated in both the RMR and the RER, then both program flags were set to one. After this step was completed, the program data was combined with the billing data by premise. The billing data was then analyzed by account to minimize noise in the model from occupancy changes during the analysis period.

There were a number of VGS customers who participated in more than one program. In the billing model, the house was considered to have received all measures installed, regardless of the program. The RMR model included all participants who installed a measure through the RMR or Home Performance components. The same criteria were used for the RLI component, as there was little or no overlap between RMR and RLI. The RER model included customers who participated only in the RER program. If a customer was a RMR participant, then any measures installed through the RER would be used to estimate the RMR measure group savings.

3.4 Non-Participant Trend Line

One aspect of billing analysis is that the difference between the pre- and post-installation energy consumption for participants only may include the effects of wider, non-program related trends in the market, in addition to any program effects. However, if non-participants are directly included in the model, there may be some confounding effects from program activities introduced into the results. For example, non-participants may install energy efficiency measures using contractors trained through VGS's program, and thus achieve higher savings.

To address these issues, a trend line reflecting the average daily consumption for non-participants was added to the regression analysis. While the trend line may still introduce some program-related effects into the models, it does not directly incorporate non-participant bills, which should mitigate the effects of program-related activities on the participant billing analysis.

VGS provided billing data for all customers. The impact evaluation team removed participants from the billing records for all customers. "Participants" were broadly defined as those customers who installed measures from 2007 through 2011 and those who received an audit but did not install measures through a VGS program. The trend line was constructed as described below:

- monthly billing records were used to determine the daily average non-participant use for every day throughout the analysis period
- the non-participant average daily ccf use for each billing cycle in the participant regression model was calculated
- the trend variable was added to the participant billing records

The trend line was added to the final models from the model selection process described above.

3.5 Model Selection

Six models were defined and the results were compared to assess whether the configuration of variables improved the model fit and also to determine whether the specific combination of variables enhanced the ability to estimate the program savings. All programs were combined in the data set used for the model selection. The candidate models are described in Table 3-2. These models were run first for the participant group only and then for the participant group with the non-participant trend line.

Table 3-2: Description of Candidate Models

Model Number	Model Description	Rank
0	Simple model; household savings	5
1	Savings by program (RMR, low income, equipment replacement)	2
2	Savings by house type (single family, multifamily units)	6
3	Heating measure groups estimated individually (envelope, heating sytem replacement, duct sealing)	3
4	Model 3 plus DHW replacement estimated individually	4
5	Model 4 plus heating system replacement by program (retrofit v. equipment replacement)	1

All of the models including the trend line ranked substantially higher than the participant only series of models. The two highest ranked models (numbers 1 and 5 in the table above) both estimated savings by program, and VGS staff also pointed out that there are substantial differences among the programs. This result led to the decision to run separate models for each program. The next ranking model with all measure groups was the selected model for the final estimation.

3.6 Regression Specifics

Following completion of a preliminary billing analysis, additional steps were conducted to review and assess the models as described at the end of this section. Given the high degree of

variability in residential energy consumption, the billing models with a higher number of homes tended to be more reliable than models with fewer homes.

Full and program-specific models were constructed. The full model included all participants with installations in the market rate and low income components of the RRP, Residential Low Income (RLI) and the Residential Equipment Replacement (RER) programs. Participants in the Home Performance with Energy Star program implemented by Efficiency Vermont were excluded from the RMR, as this program was not directly implemented by VGS. The RER program had the highest number of participants and greatly expanded the number of homes in the model. This model was used as the comparison for the candidate models. Final realization rates for each program were based on the program-specific models.

The heating degree days as calculated by VGS and included in the billing records, were incorporated into the heating-related variables. 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 period; the post-installation variable reflects the difference in heating slopes, and thus the savings. The resulting estimators were in units of ccf savings per degree day, and were multiplied by the annual heating degree days for the participants with the measure to calculate energy savings per year.

The realization rate for the homes in the model was determined by comparing the program savings from the billing models to the program-reported savings. The normalized heating degree days from 2007 through 2012 from the VGS billing data were used to estimate the final measure group and program savings. The five-year normalized heating degree days were 7,325.

The variables included in the final measure-level model are described in Table 3-3 on the next page.

Table 3-3: Variables in the Final Model

Effect Estimated	Variable Name	Interaction	Description
Customer-specific Intercept	Acct	Heating Degree Days	A separate variable for each home was included in the model to capture the average use in the house; this variable was interacted with heating degree days to allow the individual heating charcteristics of each home to be included in the model
Annual Variations in Use	year1 - year5	None	Accounts for time effects not related to the Program or other known factors for all homes in the models
Water Heater Replacement	Dhwrep	None	Dummy variable, set to 1 if the home received a water heater repair
Heating System Replacement	hsrep	Heating Degree Days	Dummy variable, set to 1 if the home received a heating system replacement; this variable was interacted with heating degree days in the post period to determine the savings per heating degree day
Envelope (Insulation and/or Air Sealing)	Env	Heating Degree Days	Dummy variable, set to 1 if the home received insulation or air sealing; this variable was interacted with heating degree days in the post period to determine the savings per heating degree day
Ducts	ducts	Heating Degree Days	Dummy variable, set to 1 if the home received duct sealing; this variable was interacted with heating degree days in the post period to determine the savings per heating degree day
None	nhdd	Daily Heating Degree Days	Accounts for weather-related changes in use
None	npcons	Non-Participant Daily Use	Trend line representing the average use for non-participants during each read cycle

Table 3-4 provides a listing of the measure groups and their associated definitions.

Table 3-4: Measure Group Definitions

Measure Group	Description
Envelope Measures*	Attic, wall and basement insulation and air sealing
Water Heater Replacement	Installation of a high efficiency water heater
Heating System Replacement	Installation of a high efficiency heating system
Ducts	Sealing and repairing ducts

* Envelope measures were most commonly installed as a group and it was not possible to estimate the savings for the specific measures since most homes had more than one of these measures installed.

3.7 Regression Diagnostics

The regression methods are based on the assumptions that the error term is independent, has a constant variance, and is normally distributed. Combining cross sectional time series (CSTS) data creates additional sources of variability. The underlying assumption behind pooling is that

the cross-sectional units are homogenous. In real applications, this is rarely the case. Energy use in homes varies widely, as does the impact of the conservation treatments. Homes with unusually high use may well have different patterns of consumption than other homes.

For ordinary least squares (OLS) regression, the assumption is that the error term is independent, has a constant variance and is normally distributed. In CSTS data sets, variation among the cross-sectional units may contribute to heteroskedasticity and the series of observations within each house may well be autocorrelated. Collinearity among the explanatory variables can also contribute to the uncertainty in the estimated intervention effects, sometimes resulting in estimators of the opposite sign.

Regression diagnostics were conducted to determine whether there were any major deviations from these assumptions.[12] The effects of heteroskedasticity (unequal variances), collinearity, autocorrelation (lack of independence among observations) and influential data points were assessed as part of the model diagnostics. A brief description of the methods is provided below, and more detail on the methods and the results of the diagnostic tests are discussed in Appendix A.

3.7.1 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.

3.7.2 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 v 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

[12] Diagnostics for RLI and RER were conducted on the final models. For RMR, the diagnostics were conducted on a model that included both RMR participants (247 homes) and EVT's HPwES program participants (26 homes).

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.

3.7.3 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.

3.7.4 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 shell 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.

3.8 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 ccf 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 ccf savings per degree day, and were multiplied by the annual heating degree days to calculate energy savings per year.

3.9 Review of the Analysis Period

All installations completed in program years 2008 through 2011 for homes with sufficient billing analysis were included in the initial model. Since it was possible that the unusually warm winter of 2011/2012 may have an impact on the results as the billing analysis relies on a strong linear relationship between natural gas use and heating degree days, the models were run with and without program year 2011 installations. The results of this comparison indicated that savings were slightly higher when the 2011 installation were excluded and reduction in the number of homes in the model did not affect the reliability of the overall results. Consequently, the final estimated savings were based on the models including installations made in program years 2008 through 2010 only to avoid the possibility that the unusual weather patterns in the winter of 2011/2012 may have introduced a downward bias into the results.

4 Results

This section provides a summary of the results of this impact evaluation by program and component. Measure level savings are presented and discussed to assist with interpreting the results. The summary of evaluated gross program savings is presented in Section 4.7.

4.1 Overview

As an overview of the program impacts, the energy savings reported by the program and the evaluated savings were compared to the pre-installation average annual energy use by program component, as shown in Table 4-1 below. Since the market rate component of the RRP is specifically targeted toward high use customers, the pre-installation average household use is the highest of the three program components, at 1,255 therms. The program reported savings indicate that the reduction in annual energy use is about 26% for participants on average, and the evaluated savings are quite similar with a 23% reduction.

Table 4-1: Comparison of Program Reported and Evaluated Savings to Annual Energy Use

	Average Pre-Installation Natural Gas Use[a]	Program Reported Savings		Evaluated Savings	
Program Component	Therms/ Year/ Household	Therms/ Year/ Household	% of Pre-Instal Use	Therms/ Year/ Household	% of Pre-Instal Use
Residential Retrofit: Market Rate	1,255	326	26%	290	23%
Residential Retrofit: Low Income	882	229	26%	143	16%
Residential Equipment Replacement	980	136	14%	127	13%

a This value reflects the average annual energy use during the pre-installation period for all homes included in the billing models.

4.2 Billing Analysis Results

The billing analysis was conducted first in its simplest form, which estimated the savings by household, and then with measure groups. The final results were based on the measure-group models, averaging the results from the participant only models and the models that included the non-participant trend line, as discussed further in Section4.4. Only homes with installations completed in 2008 through 2010 were included in the models.

Table 4-2 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-values for

envelope and heating system replacement measures are over 17 for the RMR, showing that the results are highly statistically significant for these two major measures accounting for 60% and 27% of total program reported savings, respectively.

Table 4-2: Summary of Final Regression Model Results

Program Component/ 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]
Residential Retrofit: Market Rate					
Water Heater Replacement	51	(0.407)	5.68	Therms/Day	5%
Envelope (*e.g.*, insulation, air sealing, windows and doors)	202	(0.025)	16.73	Therms/Heating Degree Day	60%
Heating System Replacement	111	(0.031)	15.93	Therms/Heating Degree Day	27%
Non-Participant Trend Variable[4]		1.236	14.59	Average Therms/Day	N/A
R-Square[5]		0.92			N/A
Total Homes in Model	247				

Program Component/ 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]
Residential Retrofit: Low Income					
Water Heater Replacement[4]	10	0.020	0.28	Therms/Day	5%
Envelope (*e.g.*, insulation, air sealing, windows and doors)	131	(0.016)	14.76	Therms/Heating Degree Day	60%
Heating System Replacement	12	(0.013)	5.36	Therms/Heating Degree Day	27%
Duct Sealing[7]	53	(0.004)	2.52	Therms/Heating Degree Day	Unknown[5]
Non-Participant Trend Variable[5]		0.765	14.67	Average Therms/Day	N/A
R-Square[6]		0.95			N/A
Total Homes in Model	132				
Residential Equipment Replacement					
Water Heater Rentals	834	(0.096)	7.99	Therms/Day	66%
Heating System Replacement	1,565	(0.014)	41.37	Therms/Heating Degree Day	34%
Non-Participant Trend Variable[5]		1.005	48.82	Average Therms/Day	N/A
R-Square[6]		0.95			N/A
Total Homes in Model	1,654				

[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. A t-value of 1.64 indicates the coefficient is statistically different from zero at the 90% confidence level.

[3] The VGS DSM Annual Reports record the savings for the RRP as a whole, *i.e.*, the market rate and low income components are combined. The percent of program savings is for the total program. Two measure categories (heat recovery and the Eco-Kit self install package), accounting for 7% of program savings, were not included in the models; thus, the percent of program savings does not add to 100%.

[4] Estimators with a t-value less than 1.645 are not statistically significant; for these two measure groups, there are too few homes in the model to be able to estimate savings.

[5] The non-participant trend variable reflects the average therms used per day by non-participants in each billing cycle.

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

[7] Duct sealing is not broken out as a separate measure in the VGS DSM Annual Report.

Some of the characteristics of the final models are described below:

- The coefficient of determination (R^2), reflecting the proportion of the change in consumption explained by the model, are quite high at 0.92, 0.95 and 0.95 for the market rate and low income residential retrofit components and the RER program, respectively.

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 40. The measure group savings were reasonably consistent under different definitions of the model variables.
- The estimators for two measure groups are not statistically significant (RLI water heater replacements and RMR duct sealing), and the water heater coefficient is also of the wrong sign (indicating extra use rather than savings). For both of these measures, the key issue is the small number of homes in the model with the measure, *i.e.*, there were insufficient homes with the measure group to be able to develop a reliable estimate of savings.
- The statistically significant measure group results are in the range of savings per project found for other, similar programs.

A review of measure-level results is presented below in Section 4.5.

4.3 Results of Regression Diagnostics

For the RMR model, six homes were identified as potential outliers. When these homes were removed from the model, the savings for the water heater replacement measure dropped by about 35%. The savings from envelope and heating system replacement measures were reasonably stable. In the original model including all participants with sufficient billing data, the realization rate for water heater replacements was 1.84; when outliers were removed, the realization rate was 0.96. The outliers are program participants and were used in the final estimates. However, this analysis points to the instability of the estimator for water heater replacements and suggests that VGS's method of estimating the savings is reasonably accurate for most homes.

The same analysis was conducted for the RLI component and the RER program, with one home removed from the RLI model and fifteen homes from the RER model. Removing the outliers had very little impact on the savings for these two models.

The RLI and RER models do not exhibit heteroskedasticity at the 5% confidence level, but the RMR model was found to be heteroskedastic. As explained above, heteroskedasticity is defined as unequal variances and could be caused by house-to-house variations in the model. However, the test for heteroskedasticity cannot necessarily separate unequal variances from misspecification of the model, i.e., the omission of key variables or inclusion of extraneous variables.

For this reason, the heteroskedasticity test was run a second time for the RMR model with the outliers removed. The results of this exercise suggest that the original high test statistic was due primarily to heteroskedasticity, as the test statistic for the model that removed the unusual homes (outliers) was substantially lower, indicating that this modified model did not exhibit heteroskedasticity.

The collinearity review led to bundling all of the insulation measures into a single variable for the RLI and RMR models, which improved our ability to estimate savings for this group of measures.

All three data sets were found to exhibit autocorrelation, as would be expected. However, autocorrelation does not introduce bias into the estimators.

4.4 Realization Rate by Program Component

The results of the participant-only and non-participant trend line models were compared to develop the final estimates of program realization rates (Table 4-3). When the non-participant trend line is added to the models, the savings are lower than with the participant only models. The magnitude of the difference varies by program. As discussed in Section 3.3, savings from the participants only model may be overstated as this model could include participant inside spillover. However, the savings with the non-participant trend line incorporate the adoption of efficiency measures outside of the program, which may be a proxy for free ridership or may include non-participant spillover effects. If the naturally occurring adoption of efficiency measures is considered a proxy for free riders, it introduces a downward bias as the primary purpose of the evaluation is to estimate gross savings impacts and the comparison group introduces these net effects. To the extent that the comparison group included non-participant spillover, the results from the non-participant model would artificially decrease savings.

In short, the participants only model may be biased upwards and the models with the non-participant trend line may be biased downward for the purposes of estimating gross savings. To try to address these issues, the final estimates were calculated by averaging the savings from the two models for each program.

Table 4-3: Comparison of Participant Only and the Non-Participant Trend Lines Models

Program	Residential Retrofit Market Rate (RMR) (Therms/year)	Residential Retrofit Low Income (RLI) (Therms/year)	Residential Equipment Replacement (RER) (Therms/year)
Participant Only Model			
Household Savings (Therms)	298	150	136
Realization Rate	0.91	0.66	1.00
Non-Participant Trend Model			
Household Savings (Therms)	283	134	117
Realization Rate	0.87	0.59	0.86
Final Estimate (Average Values)			
Household Savings (Therms)	290	142	127
Realization Rate	**0.89**	**0.62**	**0.93**
Number of Homes in the Model	247	132	1,654

The RER program has the highest realization rate at 0.93. This value represents an upper bound on the realization rate as it represents the retrofit savings rather than the market opportunity savings.

4.5 Measure Group Savings

A challenge of conducting billing analysis for natural gas consumption is to distinguish between heating and base (non-temperature dependent) use. Heating use is closely related to outdoor temperature, but the characteristics of the relationship vary from one home to the next. For this reason, house-specific heating characteristics were incorporated into the model by the inclusion of a separate heating slope for each individual home.

Experience has shown that savings from water heating measures tend to be variable in natural gas models. Consequently, the household savings are reliable, but the savings associated with specific measure groups may be variable. In addition, there were an insufficient number of homes in some measure groups for reliable estimation. With these caveat, the savings for specific measure groups are provided in Table 4-4 for context.

Table 4-4: Savings by Measure Group from the Natural Gas Billing Model

			Regression Results			
Measure Group	**Number of Homes in the Model**	**Program Savings per Home (Therms/year)**	**Evaluated Savings per Home (Therms/year)**	**Lower 90% Confidence Limit[1]**	**Upper 90% Confidence Limit[1]**	**Realization Rate**
Residential Retrofit Market Rate						
Water Heater Replacement	51	75	152	147	156	2.03
Envelope Measures	202	261	191	190	193	0.73
Heating System Replacement	111	216	228	226	230	1.05
Duct Sealing	0	0	0.0	0.0	0.0	
RRP Savings per Household	247	326	290	270	310	0.89
Residential Retrofit Low Income						
Water Heater Replacement	*10*	*67*	*-4*	*-46*	*38*	*-0.05*
Envelope Measures	131	186	125	112	138	0.67
Heating System Replacement	12	132	98	69	128	0.75
Duct Sealing	53	68	26	9	43	0.37
RLI Savings per Household	132	229	143	128	158	0.62
Residential Equipment Replacement Program						
Water Heater Replacement	834	77	41	34	48	0.53
Heating System Replacement	1,565	102	112	108	116	1.10
RER Savings per Household	1,654	136	127	122	132	0.93

Italics indicate that the savings are likely to be too small to estimate given the number of homes in the model that received the measure.
[1] The confidence intervals reflect the variability in the modeling, not the sampling precision (as no sampling was necessary).
[a] The estimator is negative, indicating extra use. However, given the few homes with this measure, it is more likely to be an indication that the model was not effective at estimating the savings.

The results for each measure group are discussed below.

4.5.1 Water Heater Replacement

For the RMR program, the realization rate for the water heater replacements is 2.03. However, when six homes identified as outliers were removed from the model, the realization rate for this measure dropped to 0.96. While the outliers are program participants and the final estimates correctly include these homes, this finding suggests that VGS is doing reasonably well at estimating savings from this measure.

For the RLI component, there were an insufficient number of homes in the model to estimate savings for DHW replacements and the savings estimate should not be considered to be reliable. This program component had the lowest number of homes of the three models, and only ten participants received a new water heater. These results should not be used to change program procedures.

The realization for the water heater replacements in the RER program is 0.53. This finding may be related to the program delivery mechanism, in which VGS offers high efficiency rental water heaters to its customers. In this situation, the characteristics of the existing water heater in use prior to the installation of the rental water heater are unknown. Given the high number of homes with this measure in the model (834) and the high statistical reliability of the regression estimator (t-value of 8.0), the estimated realization rate for this measure should be considered to be reliable as a retrofit measure. This value represents an upper bound on the realization rate as it represents the retrofit savings rather than the market opportunity savings.

Although the realization rate for high efficiency rental water heaters has high statistical reliability, this study does not address the question of whether offering standard efficiency rental water heaters is consistent with VGS's mandate to provide least cost service to VGS customers.[13] The efficiency of water heaters rented to its customers is entirely in the control of VGS. VGS is required by the Public Service Board of Vermont to utilize the principles of least cost planning in its operations. If the high efficiency water heaters are the least cost option from a societal perspective, then it may not be appropriate to be claiming savings for this measure.

4.5.2 Envelope Measures

The realization rates for both the RMR and the RLI program components are around 0.70 for envelope measures. In the RMR model with the outliers excluded, the savings from envelope measures remained reasonably consistent. The lower realization rate result may well be a reflection of the difficulty in estimating savings for these measures, as also suggested by the low realization rates for envelope measures found in two recent impact evaluations of residential programs in New York State.[14]

4.5.3 Heating System Replacements

The program estimated savings from heating system retrofits are quite accurate for the RRP and RER programs, as shown by the realization rate of 1.10 for both programs. For the RER program, this value represents an upper bound on the realization rate as it represents the retrofit savings rather than the market opportunity savings. The impact evaluations of the two residential New York programs also had a similar outcome. These results suggest that the savings from these measures are more robust and that methods of estimating program savings are producing reasonably accurate results for retrofit applications.

[13] VGS pointed out that the rental water program is administered through the non-regulated service department within Vermont Gas and VGS customers have the option to select from a list of certified contractors to supply and install their water heaters.

[14] See references, EmPower (Table ES-5) and HPwES (Table ES-2).

The realization rate for heating system replacements in the low income program is noticeably lower at 0.75. The overall realization rate for the program at 0.63 suggests that there may be issues with the methods used by the program to estimate savings.

4.5.4 Duct Sealing

The RLI program was the only one of the three with a sufficient number of homes in the model to estimate the savings from duct sealing. The realization rate for this measure was 0.37, suggesting that the current method of estimating savings used by the program is substantially overstating the savings from duct sealing.

4.5.5 Eco-Kit Self Installation

In its DSM annual reports, VGS reports that it is providing Eco-Kits to participants for them to install. The Eco-Kit includes a low flow showerhead, two faucet aerators and a coil of foam weatherstripping. VGS appears to be claiming 43 therms of annual savings for each kit sent out.[15]

There did not appear to be any Eco-Kits recorded in the program data set provided by VGS to the evaluators, and this measure was not included in the billing models. If these kits were distributed to some of the participants included in the models, these savings are likely to be included in the estimated savings for other measures.

However, it is likely that the actual impact of the Eco-Kit would be quite small. The program reported savings are high for this type of package, particularly considering that it is up to the participant to do the installation. The TRM used by Efficiency Vermont and BED allows deemed savings of 260 kWh for a low flow showerhead and 45 kWh for aerators. Converting these values to therms and adjusting for the lower efficiency of natural gas water heaters, the equivalent TRM savings for the low flow portion of VGS's Eco-Kit would be about 20 therms.

4.6 Comparison to Other Programs

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 in time are outside of the scope of the evaluation. In addition, this evaluation is also subject to uncertainty surrounding changing human behavior and economic conditions which may impact final results.

For comparison purposes, research was conducted into evaluations of similar natural gas programs as described below. These studies all contain varying kinds of uncertainty as well and

[15] The savings were estimated from the "slam dunk" end use description in VGS's DSM Annual Reports for the RIR Program, averaged over the three years of the evaluation period. VGS's annual reports also include a brochure apparently provided to customers which indicates that the Eco-Kit will save up to 86 therms per year.

are provided to offer some context relative to 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 VGS's RRP; the impact evaluation for program year 2006 found that the realization rates of 0.76 for natural gas.[16]
- Wisconsin's targeted Home Performance with Energy Star program includes the same range of measures as VGS's RRP; 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 for natural gas, respectively, when compared to the savings reported for program year 2004.[17]
- 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.[18]
- Two other NYSERDA residential programs received impact evaluations for program years 2007/2008, EmPower (residential low income) and Home Performance with Energy Star.[19] The overall natural gas realization rates for these two programs were 0.70 and 0.65, respectively.

VGS's realization rate of 0.89 +/- 6% at the 90% confidence level for the market rate component of the RRP is substantially higher than found in these other relatively recent natural gas impact evaluations.

In addition to the above natural gas impact evaluation results, the Vermont PSD recently completed an impact evaluation of Efficiency Vermont's market rate thermal efficiency Home Performance with Energy Star (HPwES) program. 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[20], and the evaluation methods were modified to address the requirements of the specific programs,[21] the VGS and HPwES realization rates are not directly comparable. The HPwES evaluation results referenced below are provided for context only. The results of the impact evaluation of Efficiency Vermont's HPwES program indicate that EVT achieved a realization rate of 0.51 +/- 13% for thermal savings and 0.86 +/- 12% for electric savings.

[16] See references, MASS, Tables 6 and 7 on pdf pages 19 and 20 (report pages 11 and 12).
[17] See references, WIS.
[18] See references, NYSERDA, Table 5-1.
[19] See references, EmPower (Table ES-1) and HPwES (Table ES-3).
[20] The VGS program uses an "operational based" rating and the EVT uses an "asset based" rating.
[21] 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.

4.7 Overview of Realization Rates

The realization rates by program and program components are shown in Table 4-4 below. Comparing the savings from the billing models to the program-reported savings indicates that about 89% of the savings estimated by VGS for the RMR component of the Residential Retrofit Program were realized, i.e., the program has a realization rate of 89%. The analysis shows that the realization rates vary by program, with the RER program having the highest realization rate of 93% and the RLI component with the lowest at 62%.

Table 4-5: Summary of Realization Rates by Program and Program Component

	Residential Retrofit: Market Rate Component	Residential Retrofit: Low Income Component	Residential Equipment Replacement
Realization Rate[1]	0.89	0.62	0.93
Lower 90% Confidence Limit[2]	0.83	0.56	0.89
Upper 90% Confidence Limit[2]	0.95	0.69	0.97
Relative Precision at 90% Confidence[2]	0.07	0.07	0.04
Number of Homes in Model	247	132	1,654

[1] The final estimates of realization rates were developed by averaging the results from the participants-only models and the models including the non-participant trend line.

[2] The confidence limits and relative precision are an indication of the variability in the model. No sampling was conducted for this analysis and all homes with sufficient billing records were included in the models. Thus, there is no sampling error associated with these results.

The estimates of precision reflect the variability in the model due to house-to-house differences in energy consumption. The realization rates were estimated at a relative precision that exceeded the 90/10 confidence/precision target.

The realization rates for the RMR component are quite good. The RMR realization rate compares favorably to the recent impact evaluation results for a similar program operated in the Northeast. It appears that VGS is making a serious and effective effort to estimating savings for its residential programs, and some tweaks to program procedures could bring the realization rates up even further.

The RER program has the highest realization rate at 0.93. This value represents an upper bound on the realization rate as it represents the retrofit savings rather than the market opportunity savings.

The lower realization for the RLI could be partially due to the specific characteristics of low income weatherization efforts. In the low income sector, participants may decide to "take back" a portion of the savings in increased comfort by raising thermostat settings or heating additional rooms. This "snap back" effect will effectively lower the savings as determined from a billing analysis. Another contributing factor to the lower realization rate could be that the methods used to estimate savings are not correctly accounting for actual conditions. The lower realization rate for the program component suggests that a review of the program procedures and methods for estimating savings should be pursued.

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. Three recommendations for the Vermont Gas programs are discussed below. The first addresses the lower realization rate found in the low-income portion of the portfolio. The other two require a reassessment of claimed savings for DHW rental tanks and Eco-Kits.

5.1.1 Work with CVOEO to improve the realization rate of the RLI

The RLI component had a realization rate of 62%. Low-income customers may be more likely to "take back" savings from efficiency programs in the form of increased comfort than other segments of the population, and this may be one reason for the lower realization rate for this program. However, it is also possible that there are other contributing factors, such as the method of estimating savings. Given that the RMR component of the program that is directly implemented by VGS has a substantially higher realization rate than the RLI component, it may be possible for VGS to offer assistance to CVOEO by reviewing savings estimation methods, installation procedures and manuals, offering training as needed or through other techniques.

5.1.2 Assess the RER water heater rental measure as part of least cost planning

During the period in this evaluation (2008 to 2010), VGS was apparently claiming savings for renting high efficiency water heaters to its customers. It could be argued that this measure is required as part of VGS's least cost planning obligations. [22] VGS should provide to the PSD an analysis demonstrating that the savings claimed for rental water heaters exceed the level of efficiency that is mandated by least cost planning. This type of analysis would remove any question as to the legitimacy of the savings claimed for this measure.

5.1.3 Review the measure characterizations for the self-installed Eco-Kits

The savings for Eco-Kits are greater than claimed for similar measures in other programs both in and outside of Vermont. Savings for showerheads and aerators and other components provided in Eco-Kits should be based on the best available information from other evaluation studies. VGS should either conduct a literature search and adopt values that are supported by evaluation studies, or adopt savings estimates that are equivalent to the estimates for electric savings in the Vermont Technical Reference Manual.

[22] VGS pointed out that the rental water program is administered through the non-regulated service department within Vermont Gas and VGS customers have the option to select from a list of certified contractors to supply and install their water heaters.

5.2 Evaluation Recommendations

Evaluation of the VGS Residential Program is an important step in insuring that all of the efficiency programs in Vermont meet generally recognized standards with respect to claimed savings. The PSD should consider conducting additional evaluation of VGS's energy efficiency portfolio using established evaluation protocols such as the *California Energy Efficiency Evaluation 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.

5.2.1 Continue with impact evaluations of other VGS efficiency programs

This evaluation of the residential retrofit programs represents a comprehensive and rigorous evaluation of VGS's residential retrofit and market opportunity programs. The C&I programs account for a substantial part of VGS's program-reported savings and they represent a part of Vermont's investment in efficiency that has not received the same level of scrutiny that is given to programs operated by Efficiency Vermont and the Burlington Electric Department.

5.2.2 Conduct studies to assess net-to-gross estimates for VGS's efficiency programs

Another potential evaluation that is indicated from this effort is a need to establish net-to-gross factors and other factors associated with VGS programs. The change in the estimated savings when the non-participant trend line is included in the models suggests that the net effects, including free riders and spillover, may be substantial. In addition, VGS is claiming savings from Eco-Kits that are installed by the participant and it is possible that some measures from the kits may not be installed. A study designed to establish how often and when these kits are installed and to develop the baseline would help establish the societal value of this initiative. Similarly a greater understanding of the rental hot water heater, with a focus on baseline and free-ridership would provide insight into the target population and effectiveness of the program.

5.3 Conclusions

This evaluation was designed to estimate gross program savings through a billing analysis of all participants with complete and reliable billing data, and both internal and external validation was conducted to ensure that the results were within a reasonable range. The internal validations consisted of a restricted billing analysis to inform the billing analysis, a review of alternative models to incorporate external influences into the billing regression and an assessment of the validity of the model. External validation consisted of comparing the results from the model to other similar programs.

Since this is a comprehensive and rigorous impact evaluation of VGS's residential retrofit and market opportunity programs, it provides the opportunity to assess the effectiveness of the programs and to identify potential problem areas that may need further investigation. For the program components directly implemented by VGS, it appears that VGS is making a serious and effective effort to estimating savings for its residential programs.

A key aspect of VGS's success with the market rate component of the RRP may be its decision to implement an operational based approach to the estimation of program savings. The Department of Energy (DOE) defines operational based rating as a rating approach that takes into account the specific characteristics of the home and lifestyle of its occupants (thermostat setting, etc.) as well as historical energy use. VGS is successfully using this strategy to improve its estimated savings.

The realization rate for the low income program delivered in partnership with CVOEO was much lower, which could be partially due to snap back among participants who decide to take their efficiency gains in the form of increased comfort. However, further investigation into other possible reasons for the lower realization rate should be pursued. It is possible that an operational based approach would be an effective tool to improve estimated savings for low income homes.

6 References

Belsley, D.A., Kuh, E., and Welsch, R.E. *Regression Diagnostics*, New York: John Wiley & Sons, Inc., 1980.

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.

Goldfeld, S.M. and R.E. Quandt. 1965. Some tests for homoscedasticity. Journal of the American Statistical Association 60 (2): 539-547.

Judge, George G., William E. Griffiths, R. Carter Hill and Tsoung-Chao Lee. 1980. *The Theory and Practice of Econometrics*. New York: John Wiley and Sons.

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 NU/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℠ Program Impact Evaluation Report, prepared for the New York State Energy Research and Development Authority. Prepared by Megdal & Associates, LLC. February 2012. (EmPower)

NYSERDA 2007-2008 Home Performance with ENERGY STAR® Program Impact Evaluation Report, prepared for the New York State Energy Research and Development Authority. Prepared by Megdal & Associates, LLC. February 2012. (HPwES)

"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 ERS, North Andover, Massachusetts. June, 2010 (NYSERDA)

Sayrs, Lois W. 1989. *Pooled time series analysis*. Newbury Park, California: Sage Publications.