# Billing Analysis of 2010 and 2011 Rooftop Unit Tune-ups

**Dan Rubado 3-12-2014**

For this analysis, we attempted to quantify the impact of rooftop HVAC unit (RTU) tune-ups on gas and electricity consumption in commercial buildings. This was a service provided through Energy Trust's Existing Buildings (EB) program from 2010-2013. RTUs are ubiquitous on small to medium size commercial buildings and they are often poorly maintained, have mechanical issues, and do not regulate the intake of outside air properly.[1] The EB program began testing in 2009 to identify maintenance and upgrade activities that could achieve significant energy savings for most RTUs and began implementing these measures through a pilot initiative in 2010. During the pilot period, measures were implemented, savings were verified, mechanical contractors were trained, and QC procedures were established. The initiative was brought to scale in 2011 and the pool of contractors qualified to deliver RTU tune-ups was substantially increased. A summary of RTU tune-ups completed over time is provided in **Table 1**. **Table 2** provides a summary of where RTU tune-ups occurred in 2010 and 2011. Three similar measures were pooled together for this analysis,[2] but actual tune-ups during 2010 and 2011 were dominated by one measure: RTU tune-ups with demand-controlled ventilation (DCV) added.

### Table 1: RTU tune-up activity by year

| Year | Measures | Sites | Claimed kWh Savings | Claimed Therm Savings |
|---|---|---|---|---|
| 2010 | 441 | 146 | 1,010,027 | 120,116 |
| 2011 | 1,481 | 607 | 4,315,872 | 572,083 |
| 2012 | 1,553 | 671 | 4,889,403 | 659,856 |

### Table 2: RTU tune-up activity by geographic region, 2010-2011

| Region | Measures | Sites | Claimed kWh Savings | % kWh Savings | Claimed Therm Savings | % Therm Savings |
|---|---|---|---|---|---|---|
| Portland Metro Area | 1,557 | 598 | 4,381,589 | 82.3% | 541,055 | 78.2% |
| Willamette Valley / North Coast | 219 | 88 | 628,098 | 1.8% | 89,891 | 13.0% |
| East of the Cascades | 132 | 54 | 299,708 | 5.6% | 60,789 | 8.8% |
| Southern Oregon / South Coast | 12 | 3 | 13,464 | 0.3% | - | 0.0% |
| Unknown | 2 | 1 | 3,040 | 0.1% | 464 | 0.1% |
| ***Total*** | ***1,922*** | ***744*** | ***5,325,899*** | **100.0%** | ***692,199*** | **100.0%** |

RTU tune-ups could only be analyzed for 2010 and 2011 due to the availability of post-treatment year consumption data. Project tracking data for these tune-ups was retrieved from Energy Trust's FastTrack database and gas and electric consumption data were retrieved from Energy Trust's utility usage database. Although most sites had both gas and electric savings from the tune-ups, gas and electric usage were analyzed separately to maximize the sample size for each fuel. To determine energy savings from RTU tune-ups, we analyzed the changes in electric and gas usage in the pre-treatment and post-treatment years at sites that received the RTU tune-ups and compared those with the changes in

[1] Robison, D., Hart, R., Price, W., and Reichmuth, H. 2008. "Field Testing of Commercial Rooftop Units Directed at Performance Verification." In *Proceedings of the 2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 3:307-318. Washington, D.C. American Council for an Energy Efficient Economy.

[2] RTU tune-up measures included the following: complete RTU package tune-up, tune-up with DCV control added, and economizer retrofit without DCV.

usage at similar commercial sites that did not receive tune-ups. The resulting “difference in differences” is the average energy savings attributable to RTU tune-ups.

### Energy Usage Data

Electric and gas utility billing data were matched to participant sites based on normalized USPS address barcodes. Many commercial sites had more than one electric or gas meter. Since we did not know which meters were associated with the treated RTUs, we aggregated all of the meters at a single site, by fuel, to obtain building-level electric and gas usage. To do this, we first normalized the meter readings to a regular interval. This was accomplished by distributing the usage associated with each meter read to the calendar months contained in the billing period, on a pro rata basis. Once the meter-level usage data were “monthified” in this way, the monthly meter-level usage at each site was summed, by fuel, to achieve monthly site usage. While combining meters undoubtedly added noise to the data, we did not have a better method for handling sites with multiple meters. Next, monthly building-level usage data were summed to create annualized electric and gas usage. Monthly data were also retained for analysis.

### Participant Sample Attrition

RTU tune-up participant sites were removed from the analysis if there were major issues with the project data or billing data, including missing information and outliers. For instance, not all sites that received tune-ups could be matched to usage data with a complete time series that encompassed the entire analysis period, including the pre-treatment, treatment, and post-treatment years. This issue caused significant attrition. In addition, sites that had large fluctuations in usage over time were dropped from the analysis. These included sites with year to year increases in consumption of more than 100% or decreases of more than 50%. Sites that had pre-treatment annual usage in the tails of the distribution, the bottom or top 2.5%, were considered outliers and were excluded from the analysis. Sites were considered outliers for a number of additional reasons and excluded, including the following:

- They received any other Energy Trust incentives or services during the analysis period;
- The total capacity of RTUs tuned-up was in the bottom or top 2.5% of the distribution;
- The number of units tuned-up was in the bottom or top 2.5% of the distribution;
- The working therm savings were in the bottom or top 2.5% of the distribution; or
- The working kWh savings were in the bottom or top 2.5% of the distribution.

Of the 146 sites that received tune-ups in 2010, 94 had complete project data and did not participate in other Energy Trust programs during the analysis period. 77 of these sites had usable gas billing data and 46 remained after the oddities and outliers had been removed, leaving 32% of sites to be analyzed. 68 of the 2010 tune-up sites had usable electric billing data and 41 remained after removing outliers, leaving 28% for analysis. Of the 571 sites that participated in 2011, 498 had complete project data and did not participate in other Energy Trust programs. 412 of these had usable gas billing data and 211 remained after the oddities and outlier had been removed, leaving 37% of sites to be analyzed. 382 of the 2011 tune-up sites had usable electric billing data and 218 remained after removing outliers, leaving 38% for analysis. The sample attrition is summarized in **Table 3.**

**Table 3: Gas and electric participant sample attrition, 2010-2011**

| Phase of Analysis | Gas Projects | | | | Electric Projects | | | |
|---|---|---|---|---|---|---|---|---|
| | 2010 | | 2011 | | 2010 | | 2011 | |
| | N | % | N | % | N | % | N | % |
| All tune-up sites with savings | 127 | 100% | 554 | 100% | 129 | 100% | 543 | 100% |
| Sites with usable address | 94 | 74% | 494 | 89% | 84 | 65% | 482 | 89% |
| Sites matched to billing data | 77 | 61% | 412 | 74% | 68 | 53% | 382 | 70% |
| Sites removed from analysis: | -32 | -25% | -206 | -37% | -30 | -23% | -167 | -31% |
| Installed another measure during analysis period | -21 | -17% | -101 | -18% | -19 | -15% | -104 | -19% |
| Pre-treatment usage outlier | -4 | -3% | -20 | -4% | -7 | -5% | -14 | -3% |
| *Lowest 2.5% pre-treatment usage* | *-2* | -2% | *-10* | -2% | *-3* | -2% | *-8* | -1% |
| *Highest 2.5% pre-treatment usage* | *-2* | -2% | *-10* | -2% | *-4* | -3% | *-6* | -1% |
| Large pre-to-post change in usage (+100% / -50%) | -9 | -7% | -69 | -12% | -4 | -3% | -40 | -7% |
| Missing usage data during analysis period | -3 | -2% | -50 | -9% | -3 | -2% | -36 | -7% |
| RTU capacity outlier | -3 | -2% | -14 | -3% | -2 | -2% | -14 | -3% |
| *Lowest 2.5% RTU capacity* | *-3* | -2% | *-3* | -1% | *-2* | -2% | *-3* | -1% |
| *Highest 2.5% RTU capacity* | *0* | 0% | *-11* | -2% | *0* | 0% | *-11* | -2% |
| RTU quantity outlier | -1 | -1% | -11 | -2% | -1 | -1% | -9 | -2% |
| *Lowest 2.5% RTU quantity* | *0* | 0% | *0* | 0% | *0* | 0% | *0* | 0% |
| *Highest 2.5% RTU quantity* | *-1* | -1% | *-11* | -2% | *-1* | -1% | *-9* | -2% |
| Claimed savings outlier | -3 | -2% | -14 | -3% | 0 | 0% | -12 | -2% |
| *Lowest 2.5% estimated savings* | *-2* | -2% | *-3* | -1% | *0* | 0% | *-1* | 0% |
| *Highest 2.5% estimated savings* | *-1* | -1% | *-11* | -2% | *0* | 0% | *-11* | -2% |
| **Total sites available for analysis** | **45** | **35%** | **206** | **37%** | **38** | **29%** | **215** | **40%** |

### Description of Participant Sample

The 2010 gas sample had an average of 2.3 RTUs tuned-up per site with a total capacity of 16.3 tons. The average gas savings claimed by the EB program for these tune-ups was 894 therms, or 14% of pre-treatment annual usage. The 2010 electric sample had an average of 2.3 RTUs tuned-up with a total capacity of 16.1 tons. The average electric savings claimed was 4,641 kWh, or 1.6% of pre-treatment annual usage. In 2011, the gas sample had an average of 2.6 RTUs tuned-up with a total capacity of 16.3 tons. The average gas savings was 925 therms, or 18%. The 2011 electric sample had an average of 2.3 RTU tune-ups per site with a total capacity of 14.6 tons. The average electric savings claimed was 5,707 kWh, or 2.7%. Pre-treatment annual gas and electric usage dropped substantially from 2010 to 2011, potentially indicating that smaller sites were targeted in 2011. Contrary to the trend to lower pre-treatment usage, the total energy savings claimed per site and the percent savings went up from 2010 to 2011. This is likely due to changes that were made to the tune-up measures in mid-2011, including changes to the deemed savings per ton. The gas and electric sample site characteristics for 2010 and 2011 are detailed in **Table 4**.

**Table 4: Gas and electric participant sample, mean site characteristics, 2010 and 2011**

| Year | Fuel | Tune-up Site N | Pre-Treatment Annual Usage | Average # of RTUs Tuned-up | Capacity of RTUs Tuned-up (tons) | Average Savings Claimed | % Savings Claimed |
|---|---|---|---|---|---|---|---|
| 2010 | Gas (therms) | 45 | 6,357 | 2.4 | 16.6 | 906 | 14.3% |
| | Electric (kWh) | 38 | 303,238 | 2.4 | 16.6 | 4,776 | 1.6% |
| 2011 | Gas (therms) | 206 | 5,181 | 2.6 | 20.2 | 978 | 18.9% |
| | Electric (kWh) | 215 | 213,467 | 2.4 | 17.7 | 5,750 | 2.7% |

### Comparison Group Selection

To control for non-programmatic trends in energy usage and external events, billing analyses such as this must include a comparison group. For both the gas and electric analyses, two comparison groups were constructed—a stratified random sample of non-participant commercial sites and a sample of future participants. Non-participants were defined as sites at a unique address with at least one commercial utility account that did not participate in an Energy Trust program during the analysis period. Future participants were sites that had an RTU tune-up after the analysis period and did not participate in any other Energy Trust programs during the analysis period.

Future participant sites are believed to be more comparable to the participant group because they have to meet the same requirements for participation and they are known commercial buildings with at least one RTU on the roof. Furthermore, it is likely that the future participant sites more closely resemble the participant sites and that they have similar RTUs as well. However, there are often insufficient numbers of future participants to construct a viable comparison group and in those cases it is useful to use non-participants for the comparison group. For the 2010 analysis, we were able to create both non-participant and future participant comparison groups for gas and electric usage. For the 2011 analysis, there were insufficient future RTU tune-up participants (those that participated in 2013) to create comparison groups for gas or electric, so we relied on the non-participant comparison group in that year. However, the impact of the future participant comparison group on the results can be assessed in the 2010 analysis, where both comparison groups were used. Using the non-participant comparison group resulted in slightly more conservative savings estimates in 2010.

The non-participant comparison groups were stratified based on pre-treatment annual usage. To determine the usage categories, the participant sites were grouped into annual usage quintiles. An equal number of non-participant sites were then randomly selected from each category of annual usage, creating a comparison group that resembled the participant group in its distribution of annual energy usage. Due to sample size limitations with the future participant group, they were not matched to the usage quintiles of the participant group. **Tables 5 and 6** illustrate that all three groups had comparable gas and electric consumption in the pre-treatment period.

**Table 5: Raw annual energy usage for participant and comparison sites, 2010**

| Fuel | Analysis Group | N | Pre-Treatment Annual Usage (2009) | Post-Treatment Annual Usage (2011) | Change in Usage |
|---|---|---|---|---|---|
| Gas (therms) | Participant | 45 | 6,357 | 6,147 | -210 |
| | Non-participant | 2,500 | 6,131 | 6,166 | +35 |
| | Future Participant | 246 | 6,724 | 6,834 | +110 |
| Electric (kWh) | Participant | 38 | 303,238 | 289,262 | -13,976 |
| | Non-participant | 2,500 | 296,187 | 289,797 | -6,390 |
| | Future Participant | 200 | 243,342 | 235,715 | -7,627 |

**Table 6: Raw annual energy usage for participant and comparison sites, 2011**

| Fuel | Analysis Group | N | Pre-Treatment Annual Usage (2010) | Post-Treatment Annual Usage (2012) | Change in Usage |
|---|---|---|---|---|---|
| Gas (therms) | Participant | 206 | 5,181 | 5,552 | +371 |
| | Non-participant | 2,500 | 4,945 | 5,081 | +136 |
| Electric (kWh) | Participant | 215 | 213,467 | 220,107 | +6,640 |
| | Non-participant | 2,500 | 212,912 | 214,160 | +1,248 |

### Billing Analysis Methodology

The billing analysis was done separately for each treatment year and fuel type, so 2010 treatment and comparison sites were analyzed separately from 2011 treatment and comparison sites. Gas usage was also analyzed separately from electric usage, although a majority of treated sites were used in both the gas and electric analyses. Monthly energy usage data and building characteristics data for all treatment and comparison sites were organized into a longitudinal, monthly dataset. The dataset contained 24 data points for each site, with 12 months in the pre-treatment year and 12 months in the post-treatment year. Each site in the dataset was matched to the nearest weather station within the same climate zone. For each station, we calculated the monthly heating degree-days (HDD) and cooling-degree days (CDD), with a base temperature of 65°F. Monthly flags were created indicating whether each record in the series occurred in the pre-treatment or post-treatment year. Flags were also assigned to indicate whether a site was part of the RTU tune-up treatment group, the future participant comparison group, or the non-participant comparison group.

Once the longitudinal dataset was prepared, we conducted a first level of regression analysis in Stata/SE v12.1 (StataCorp LP, College Station, TX) to determine the overall gas and electric savings for 2010 and 2011 RTU tune-up sites. For each analysis year, fuel type, and comparison group, we created a linear mixed effects regression model (using the xtmixed procedure in Stata), treating the monthly usage as a repeated measurement on each site. A restricted maximum likelihood method was selected to calculate the parameter estimates. Monthly usage was modeled as a function of monthly HDD and CDD, the pre/post-treatment year flag, the treatment /comparison group flag, and a random intercept for each. An interaction term was added to model the effect of the treatment year between the treatment and comparison groups—the difference in differences in energy usage. The following equation describes the model:

$$Usage_{ij} = (\beta_0 + u_i) + \beta_1 HDD_{ij} + \beta_2 CDD_{ij} + \beta_3 TxGroup_i + \beta_4 PostTx_j + \beta_5 TxGroup_i * PostTx_j + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma^2),\ u_i \sim N(0, \sigma^2)$$

Where:

$Usage_{ij}$ is the total electric or gas usage for site $i$ during month $j$,
$\beta_0$ is the fixed intercept for all sites,
$HDD_{ij}$ is the Heating Degree-Days for site $i$ during month $j$,
$CDD_{ij}$ is the Cooling Degree-Days for site $i$ during month $j$,
$TxGroup_i$ $\{0,1\}$ is a dummy variable where 1 indicates that site $i$ is part of the treatment group,
$PostTx_j$ $\{0,1\}$ is a dummy variable where 1 indicates that month $j$ occurred after the treatment period,
$\epsilon_{ij}$ is the model error for site $i$ during month $j$, and,
$u_i$ is the random intercept for site $i$ and is independent from $\epsilon_{ij}$.

This model allowed us to estimate the monthly energy savings attributable to RTU tune-ups, while controlling for the effects of weather. The model was run separately using the future participant and non-participant comparison groups to obtain estimates of monthly energy savings using each. The model was also run separately for each fuel type and treatment year to estimate gas and electric savings for both 2010 and 2011 tune-ups. The average monthly savings estimates were multiplied by 12 to calculate the annual energy savings.

### Overall RTU Tune-up Savings Results

**Tables 7 and 8** show overall weather normalized gas and electric savings results for RTU tune-ups done in 2010 and 2011, respectively. In 2010, the weather normalized gas savings were 6% of pre-treatment consumption, well below the average savings claimed for RTU tune-ups in 2010 (14%) and not statistically different from zero. Weather normalized electric savings in 2010 were 2.0 to 2.8%, which was higher than the average savings claimed in 2010 (1.6%); however, the weather normalized electric savings were only significant when compared to future participants.

In 2011, the results were reversed, with significant increases in both gas and electric usage at sites that received RTU tune-ups. The 2011 savings were -5% for gas and -2.4% for electric with realization rates below zero. To explore the potential reasons for the low 2010 savings and the anomalous 2011 results, we conducted further analysis.

### Table 7: Weather normalized gas and electric savings estimates for RTU tune-ups, 2010

| Fuel | Comparison Group | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|
| Gas (therms) | Future Participant | 382 (-46, 810) | 6.0% | 42% | 0.142 |
| | Non-participant | 383 (-106, 871) | 6.0% | 42% | 0.198 |
| Electric (kWh) | Future Participant | 6,072 (232, 11,913) | 2.0% | 127% | 0.087 |
| | Non-participant | 8,558 (-835, 17,952) | 2.8% | 179% | 0.134 |

**Table 8: Weather normalized gas and electric savings estimates for RTU tune-ups, 2011**

| Fuel | Comparison Group | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|
| Gas (therms) | Non-participant | -259 (-449, -69) | -5.0% | -26% | 0.025 |
| Electric (kWh) | Non-participant | -5,226 (-8,622, -1,830) | -2.4% | -91% | 0.011 |

### Sensitivity Analysis on the Influence of Weather

As a check on the analysis, we calculated non-weather normalized savings for each fuel type and treatment year from raw annualized, or unadjusted, energy usage (**Tables 9 and 10**), using a simple difference-in-differences calculation. Unadjusted savings were calculated as the difference in the change in usage (pre-treatment to post-treatment year) between the treatment and comparison group. We also checked these estimates against the output from a simplified version of the mixed effects regression model presented above, with the HDD and CDD terms removed. This model followed an equation of the form: $Usage_{ij} = (\beta_0 + u_i) + \beta_1 TxGroup_i + \beta_2 PostTx_j + \beta_3 TxGroup_i * PostTx_j + \epsilon_{ij}$. The savings estimates derived from this regression model and the hand calculated difference-in-differences were identical.

The variation in average annual weather from the pre-treatment to post-treatment year for the 2010 treatment sites was relatively small in terms of HDD. Differences in HDD for the 2011 treatment year were essentially zero. For both treatment years, more significant variations in CDD occurred, but the overall frequency of warm weather was minor compared to the occurrence of cold weather. For gas and electric usage, the non-weather normalized savings estimates were very close to the regression-derived, weather normalized results. The similarity in results between the unadjusted and weather normalized savings held true for all sub-category analyses as well. Thus, we are confident that we can rule out weather or weather normalization errors as contributors to the unexpected 2011 savings results.

**Table 9: Unadjusted gas and electric savings estimates for RTU tune-ups, 2010**

| Fuel | Comparison Group | Unadjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|
| Gas (therms) | Future Participant | 321 (-253, 894) | 5.0% | 35% | 0.358 |
| | Non-participant | 246 (-331, 822) | 3.9% | 27% | 0.484 |
| Electric (kWh) | Future Participant | 6,349 (243, 12,455) | 2.1% | 133% | 0.087 |
| | Non-participant | 7,586 (-1,878, 17,051) | 2.5% | 159% | 0.187 |

**Table 10: Unadjusted gas and electric savings estimates for RTU tune-ups, 2011**

| Fuel | Comparison Group | Unadjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|
| Gas (therms) | Non-participant | -235 (-457, -13) | -4.5% | -24% | 0.082 |
| Electric (kWh) | Non-participant | -5,391 (-8,803, -1,979) | -2.5% | -94% | 0.009 |

### Analysis of RTU Capacity Tuned-up

The next level of analysis was to determine if we could identify differences in the magnitude of savings for different categories of RTU capacity. We created four categories of RTU capacity and re-ran the regression models for each fuel, using only the subset of sites in each category. The categories represented projects with total RTU capacity tuned-up of five tons, 6-9 tons, 10-19 tons and 20 or more tons. For consistency, all RTU project categories were analyzed against the non-participant comparison group for the specified fuel type and treatment year. The sample size of sites was small for some sub-groups and the resulting savings estimates for some RTU capacity categories were not very precise, particularly for 2010 tune-ups. However, the results were useful in identifying potential trends in the impact of RTU capacity tuned-up on savings.

**Tables 11 and 12** summarize the results of the savings analysis by capacity of RTUs tuned-up in 2010 and 2011, respectively. We would expect that projects with more RTU capacity tuned-up would have higher savings, but this was not the trend observed. In 2010, the sample sizes of tune-up sites for individual RTU capacity categories were very small and the savings estimates were not statistically different from zero. However, as a directional indicator, the results show that savings were positive and relatively robust in 2010 across all categories except for the largest capacity category, where gas and electric savings were much lower. We were not able to detect any significant differences between categories in the 2010 results.

In 2011, there were much larger sample sizes of treated buildings for each RTU capacity category. However, even with larger samples, most of the savings estimates were not statistically different from zero. As a directional indicator, both 2011 gas and electric savings were negative for three of four capacity categories, indicating increases in usage, on average. The 10-19 ton category had relatively large negative savings for both gas and electric that were highly statistically significant (-11% for gas and -10% for electric). For 2011 electric savings, all categories were very close to zero except for the 10-19 ton category, which had relatively large negative savings (-9.2%) that were highly significant.

**Table 11: Weather normalized gas and electric savings estimates by capacity of RTUs tuned-up, using the non-participant comparison group, 2010**

| Fuel | RTU Capacity (tons) | N | Average # of RTUs | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|---|---|
| Gas (therms) | 5 | 6 | 1.0 | 853<br>(-480, 2,186) | 15.9% | 283% | 0.293 |
| | 6-9 | 5 | 1.2 | 1,314<br>(-147, 2,775) | 15.0% | 287% | 0.139 |
| | 10-19 | 19 | 2.2 | 204<br>(-546, 954) | 3.6% | 27% | 0.654 |
| | ≥20 | 15 | 3.5 | 109<br>(-735, 953) | 1.6% | 7.4% | 0.831 |
| Electric (kWh) | 5 | 6 | 1.0 | 2,673<br>(-20,890, 26,236) | 2.2% | 162% | 0.852 |
| | 6-9 | 5 | 1.0 | 3,760<br>(-22,055, 29,574) | 1.2% | 196% | 0.811 |
| | 10-19 | 15 | 2.1 | 13,134<br>(-1,785, 28,052) | 4.8% | 389% | 0.148 |
| | ≥20 | 12 | 4.1 | 7,786<br>(-8,909, 24,480) | 1.9% | 84% | 0.443 |

**Table 12: Weather normalized gas and electric savings estimates by capacity of RTUs tuned-up per site, using the non-participant comparison group, 2011**

| Fuel | RTU Capacity (tons) | N | Average # of RTUs | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|---|---|
| Gas (therms) | 5 | 38 | 1.0 | -98<br>(-526, 329) | -2.8% | -39% | 0.705 |
| | 6-9 | 34 | 1.0 | 61<br>(-386, 509) | 2.3% | 19% | 0.822 |
| | 10-19 | 57 | 2.0 | -715<br>(-1,064, -366) | -11.0% | -115% | 0.001 |
| | ≥20 | 77 | 4.6 | -141<br>(-443, 161) | -2.3% | -7.5% | 0.442 |
| Electric (kWh) | 5 | 39 | 1.0 | 1,964<br>(-5,706, 9,633) | 1.6% | 122% | 0.674 |
| | 6-9 | 41 | 1.0 | -3,099<br>(-10,578, 4,380) | -2.6% | -141% | 0.495 |
| | 10-19 | 66 | 1.9 | -18,199<br>(-24,253, -12,146) | -9.6% | -439% | <0.001 |
| | ≥20 | 69 | 4.4 | 1,830<br>(-3,977, 7,638) | 0.5% | 16% | 0.604 |

**Analysis of Number of Units Tuned-up**

To investigate the potential effects of project size and quality assurance practices on savings per site, we analyzed savings by number of RTUs tuned-up at each site. We created two categories of project size based on the number of units tuned-up: projects with fewer than four RTU tune-ups and those with four or more. These categories were based on the EB program's quality assurance threshold, where 100% of projects with four or more tune-ups received inspections. We then re-ran the regression models for each fuel and treatment year, using only the subset of sites in each project size category. For consistency, all RTU project sites were analyzed against the non-participant comparison group for the specified fuel type and treatment year.

An additional regression model was specified, using just the treatment group sites, to estimate the difference in savings between the two categories of number of units for each fuel type and treatment year. This model followed an equation of the form: $Usage_{ij} = (\beta_0 + u_i) + \beta_1 HDD_{ij} + \beta_2 CDD_{ij} + \beta_3 UnitsCat_i + \beta_4 PostTx_j + \beta_5 UnitsCat_i * PostTx_j + \epsilon_{ij}$, where $UnitsCat_i$ is a dummy variable where 1 indicates the category of projects with four or more RTU tune-ups. $\beta_5$, the coefficient of the interaction term ($UnitsCat_i * PostTx_j$), represents the weather normalized difference in monthly savings between the two categories. This coefficient was multiplied by 12 to achieve an estimate of the annual difference in energy savings.

**Tables 13 and 14** summarize the results of our analysis of tune-up savings by number of units tuned-up for 2010 and 2011, respectively. In 2010, sample size of projects with four or more RTUs tuned-up was very small, so it was difficult to precisely determine the savings for this category and determine if they were different from the smaller category. However, there was significantly more electric savings in 2010 at sites with fewer than four tune-ups, counter to what we would expect, but similar to our findings by RTU capacity. This is probably a tenuous difference, given the small sample size of the four or more tune-ups category.

In 2011, the sample sizes were more robust and there were negative savings across the board. The negative savings were relatively large and statistically significant in the fewer than four tune-ups category for both gas and electric, while savings in the four or more tune-ups category were close to zero. The differences in savings between project size categories in 2011 were not statistically significant.

**Table 13: Weather normalized gas and electric savings estimates by number of RTUs tuned-up per site, using the non-participant comparison group, 2010**

| Fuel | # of RTUs | N | Average RTU Capacity (tons) | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings | Savings Difference | p-value of Difference |
|---|---|---|---|---|---|---|---|---|---|
| Gas (therms) | <4 | 39 | 14.0 | 446 (-79, 970) | 7.9% | 56% | 0.162 | 490* | 0.403 |
| | ≥4 | 6 | 33.3 | -28 (-1,362, 1,305) | -0.3% | -1.7% | 0.972 | Ref** | Ref** |
| Electric (kWh) | <4 | 31 | 11.9 | 12,244 (1,850, 22,639) | 4.6% | 369% | 0.053 | 20,201* | 0.035 |
| | ≥4 | 7 | 37.5 | -7,761 (-29,587, 14,065) | -1.7% | -69% | 0.559 | Ref** | Ref** |

* The weather normalized, modeled mean difference in savings between the two categories of RTUs tuned-up per site, where ≥4 RTUs is the reference category.

**Indicates the reference category for the comparison of mean savings.

**Table 14: Weather normalized gas and electric savings estimates by number of RTUs tuned-up per site, using the non-participant comparison group, 2011**

| Fuel | # of RTUs | N | Average RTU Capacity (tons) | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings | Savings Difference* | p-value of Difference |
|---|---|---|---|---|---|---|---|---|---|
| Gas (therms) | <4 | 151 | 11.2 | -294 (-513, -76) | -6.3% | -54% | 0.027 | -134 | 0.620* |
| | ≥4 | 55 | 45.0 | -161 (-517, 195) | -2.4% | -7.5% | 0.457 | Ref** | Ref** |
| Electric (kWh) | <4 | 168 | 11.3 | -6,079 (-9,896, -2,262) | -3.5% | -169% | 0.009 | -3,763 | 0.429* |
| | ≥4 | 47 | 40.9 | -2,208 (-9,227, 4,812) | -0.6% | -16% | 0.605 | Ref** | Ref** |

* The weather normalized, modeled mean difference in savings per site between the two categories of number of RTU tune-ups, where ≥4 RTUs is the reference category.
**Indicates the reference category for the comparison of mean savings.

### Analysis of RTU Contractors

Based on what we know about the expertise required to properly tune-up RTUs, the findings of other studies on RTU tune-ups, and Energy Trust's past experience, it is plausible that RTU tune-up savings could vary depending on the HVAC contractor performing the tune-ups. Differences in savings between firms could be due to a number of factors including site selection, selection of RTUs to be tuned-up, baseline conditions, or quality of work. To determine the impact that contractors may have had on RTU tune-up savings, we selected the top four contractors from 2010 and 2011 and analyzed each separately. These firms were designated 'Z', 'Y', 'X' and 'W' to keep the identity of each firm confidential. Regression models were re-run for each fuel and treatment year using only the subset of sites where each contractor had performed tune-ups. We also re-ran each model using the inverse subset of sites for each contractor. For consistency, all RTU project sites were analyzed against the non-participant comparison group for the specified fuel type and treatment year.

An additional regression model was specified for each contractor, using just the treatment group sites, to estimate the difference in savings between that contractor and all others for each fuel type and treatment year. This model followed an equation of the form: $Usage_{ij} = (\beta_0 + u_i) + \beta_1 HDD_{ij} + \beta_2 CDD_{ij} + \beta_3 Contractor_i + \beta_4 PostTx_j + \beta_5 UnitsCat_i * PostTx_j + \epsilon_{ij}$, where $Contractor_i$ is a dummy variable where 1 indicates sites that were treated by the contractor of interest. $\beta_5$, the coefficient of the interaction term ($Contractor_i * PostTx_j$), represents the weather normalized difference in monthly savings between the contractor of interest and all other firms. This coefficient was multiplied by 12 to achieve an estimate of the annual difference in energy savings.

The results of the analysis by contractor are summarized in **Tables 15 and 16.** They show that there were large variations in savings by HVAC contractor in both 2010 and 2011. Contractors 'Y' and 'X' were consistently associated with positive savings across the board in both years, although these were based on small samples and were not always significantly different from zero. Contractor 'Z' did not participate in 2010, but was responsible for roughly half of the tune-ups done in 2011. Unfortunately, in 2011, contractor 'Z' was significantly associated with large, negative savings for both fuels. Gas and electric savings for contractor 'Z' were much lower than all other firms combined. In fact, when we removed sites tuned-up by contractor 'Z' from the analysis, 2011 gas savings reverted to roughly zero and electric savings became slightly, but not significantly, positive (**Table 17**). Contractor 'W' was unusual in that they were significantly associated with large, negative savings in 2010, but reversed to positive savings, albeit insignificant, in 2011. The 2010 sample size for contractor 'W' was very small, so the 2010 anomaly may not be indicative of a real difference.

**Table 15: Weather normalized gas and electric savings estimates for top HVAC contractors, using the non-participant comparison group, 2010**

| Fuel | Firm | N | Average RTU Capacity (tons) | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings | Savings Difference With All Other Firms* | p-value of Difference |
|---|---|---|---|---|---|---|---|---|---|
| Gas (therms) | Z | 0 | - | - | - | - | - | - | - |
| | Y | 20 | 19.1 | 431 (-300, 1,162) | 9.2% | 46% | 0.332 | 57 | 0.887 |
| | X | 14 | 15.8 | 676 (-198, 1,550) | 7.3% | 76% | 0.204 | 487 | 0.257 |
| | W | 5 | 13.4 | -795 (-2,256, 666) | -11.1% | -88% | 0.371 | -1,332 | 0.035 |
| Electric (kWh) | Z | 0 | - | - | - | - | - | - | - |
| | Y | 13 | 17.2 | 21,398 (5,370, 37,425) | 5.4% | 414% | 0.028 | 19,288 | 0.013 |
| | X | 13 | 13.8 | 12,742 (-3,295, 28,779) | 3.9% | 265% | 0.191 | 5,873 | 0.453 |
| | W | 7 | 23.6 | -11,707 (-33,524, 10,111) | -4.9% | -225% | 0.377 | -24,842 | 0.009 |

* The weather normalized, modeled mean difference in savings per site between a given firm and all other firms, where all other firms is the reference category.

**Table 16: Weather normalized gas and electric savings estimates for top HVAC contractors, using the non-participant comparison group, 2011**

| Fuel | Firm | N | Average RTU Capacity (tons) | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings | Savings Difference With All Other Firms* | p-value of Difference |
|---|---|---|---|---|---|---|---|---|---|
| Gas (therms) | Z | 112 | 20.7 | -456 (-708, -204) | -8.6% | -47% | 0.003 | -436 | 0.069 |
| | Y | 12 | 35.2 | 335 (-416, 1,086) | 11.1% | 21% | 0.463 | 642 | 0.209 |
| | X | 8 | 20.4 | 957 (22, 1,891) | 6.3% | 115% | 0.092 | 1,263 | 0.041 |
| | W | 18 | 17.0 | 183 (-430, 797) | 6.1% | 17% | 0.624 | 482 | 0.255 |
| Electric (kWh) | Z | 125 | 19.2 | -11,990 (-16,410, -7,571) | -5.5% | -183% | <0.001 | -16,056 | <0.001 |
| | Y | 3 | 24.5 | 28,266 (659, 55,873) | 7.5% | 380% | 0.092 | 33,345 | 0.046 |
| | X | 6 | 12.5 | 6,763 (-12,820, 26,347) | 1.2% | 187% | 0.570 | 15,451 | 0.297 |
| | W | 17 | 13.0 | 870 (-10,735, 12,475) | 0.5% | 28% | 0.902 | 6,736 | 0.355 |

* The weather normalized, modeled mean difference in savings per site between a given firm and all other firms, where all other firms is the reference category.

**Table 17: Weather normalized gas and electric savings estimates for RTU tune-ups, excluding sites that received a tune-up from contractor 'Z', using the non-participant comparison group, 2011**

| Fuel | N | Average RTU Capacity (tons) | Adjusted Savings Per Site (90% CI) | % Savings | Realization Rate | p-value of Savings |
|---|---|---|---|---|---|---|
| Gas (therms) | 94 | 19.7 | -23 (-298, 252) | -0.5% | -2.4% | 0.889 |
| Electric (kWh) | 90 | 15.7 | 4,163 (-916, 9,243) | 2.0% | 89% | 0.178 |

## Conclusions

RTU tune-ups conducted through the EB program had much lower than expected gas savings per site in 2010 and were even associated with significant increases in gas and electric usage in 2011. Further analysis showed no discernible pattern in savings by RTU capacity tuned-up or number of RTUs tuned-up that could help explain this unexpected swing in savings. The savings appear to vary substantially by HVAC contractor. Unfortunately, contractor 'Z' did roughly half of the tune-ups in 2011 and appears to have been a significant driver of the observed negative savings. However, even when we factored out contractor 'Z', the 2011 savings numbers still hovered close to zero, so clearly it was not just a single contractor that swayed the results.

It appears that there were major problems with the EB program's RTU tune-up initiative, especially in 2011. Although some savings were achieved in the 2010 pilot year, those savings were erased in 2011. This coincides with the year that the initiative was expanded, new contractors joined the pool and quality control inspections were reduced. The HVAC contractors performing the tune-ups appear to be somehow involved in this decline in savings. This could be related to contractors' selection of buildings, selection of RTUs to be tuned-up, or by bringing under-ventilated buildings up to code by improving ventilation. It could also be a matter of poor work quality, insufficient technician training, ineffective tune-up practices or relaxed quality control procedures. Further investigation is warranted to verify these findings, explore the potential causes for the failure of this measure to achieve the expected energy savings, and determine if and how a modified tune-up initiative could capture reliable, cost-effective savings. Consequently, Energy Trust has commissioned an impact evaluation of 2012 RTU tune-ups to be completed in 2014, which will involve interviews with HVAC contractors and site visits to a sample of sites that received tune-ups.