Tracking Small Business Program Performance with Annual Billing Analyses

Andrew Nawrocki, Navigant, Boulder, CO
Brian Billing, AEP Ohio, Gahanna, OH
Randy Gunn, Navigant, Chicago, IL
Derya Eryilmaz, Navigant, Houston, TX

ABSTRACT

This paper explores how a responsive Small Business Energy Efficiency Program can benefit from targeted evaluation work, underpinned by an annual billing analysis serving as the primary gross savings estimation technique. The Small Business "Express" Program works exclusively with an Implementation Contractor (IC) that is responsible for outreach, auditing, coordinating installations, payment processing and reporting to the utility. The program has evolved since inception in 2011 and has undergone a change in IC, numerous updates to the reported savings methodology, and measure mix.

Because a billing analysis of the entire population of participants inherently captures all changes in energy consumption at the facility level, care must be taken to ensure that the model controls for appropriate factors, including facility energy consumption, project size, seasonality, and other factors. The billing analysis employed by the authors use participants as their own controls for savings estimates by employing a variation in adoption model. One limitation of the billing analysis is that it does not shed insight into the drivers of program realization rates; therefore, additional primary and secondary research informed these drivers and were used to explain the results. These findings, corroborated by overarching realization rates from the annual billing analysis, have been applied by the utility and IC to improve realization rates over time. This longitudinal view shows that performance can be adequately tracked and improved over time via a billing analysis.

Introduction

AEP Ohio provides a suite of energy efficiency programs designed to help both residential and non-residential customers save energy. The authors provide independent evaluation services to AEP Ohio to both verify gross savings estimates and provide meaningful recommendations to improve the performance of energy efficiency programs. This paper focuses on tracking the performance of a single business energy efficiency program specifically tailored to small business customers from 2011 to 2016.

This program provides a one-stop, turnkey service to small businesses for energy efficient lighting and refrigeration equipment upgrades. Eligible customers have an annual consumption of 200,000 kWh or less, or a maximum of 100kW demand service. Savings estimates are based on prescriptive formulas for simplicity and auditability, while tailoring key parameters, such as hours of use, on a fixture-by-fixture basis. The implementation contractor serves as the contact point for the program to simplify the participation process for small businesses with limited resources and energy efficiency expertise.

The following timeline shows the key program milestones from 2011 through 2016. The remainder of this section walks through the specific program adjustments implemented since 2011 and their relevance to the parameters used to calculate energy savings estimates.

- 2011
  - Program launch; program ramped up throughout the year.
  - Program only incented lighting fixtures, primarily T8 fluorescent retrofits.
  - Design allowed independent trade allies to submit projects, with limited initial oversight.

- 2012
  - Program suspended mid-year while new implementation contractor was selected.
- 2013
  - Program relaunched with new implementation contractor
  - Program design centralized program management and operation.
  - Included provision to account for lighting burnouts in existing fixtures.
- 2014
  - Added refrigeration measures, primarily LED case lighting, ECM motors and anti-sweat heater controls.
  - Implemented auditor training to ensure consistency between projects, and installer training to ensure that changes to equipment orders were reflected in the tracking data.
- 2015
  - Shifted lighting measures primarily to LED retrofit fixtures.
  - Continued to refine auditor and installation contractor training, e.g. consistent hours of use estimates and installation contractor training to ensure change orders are reflected in ex ante tracking data.
  - Continued to enhance quality control, e.g. enhancing the auditor table-based tool to check for unrealistic savings estimates in real time.
- 2016
  - Updated baseline wattage assumptions.
  - Continued to refine auditor and installation contractor training.
  - Continued to enhance quality control.

Participation varied between 2011 and 2016, but was approximately 400-1000 unique projects per year as shown in Figure 1 below. Program size was driven by many factors, including program goals, budget, marketing efforts, and market forces. Program reported energy savings are also shown in the figure, and have historically averaged between 7,000 MWh and 13,000 MWh. With the presence of historical data and relatively high (greater than 10%) savings per participant, the authors determine that this program was suitable for econometric modelling.
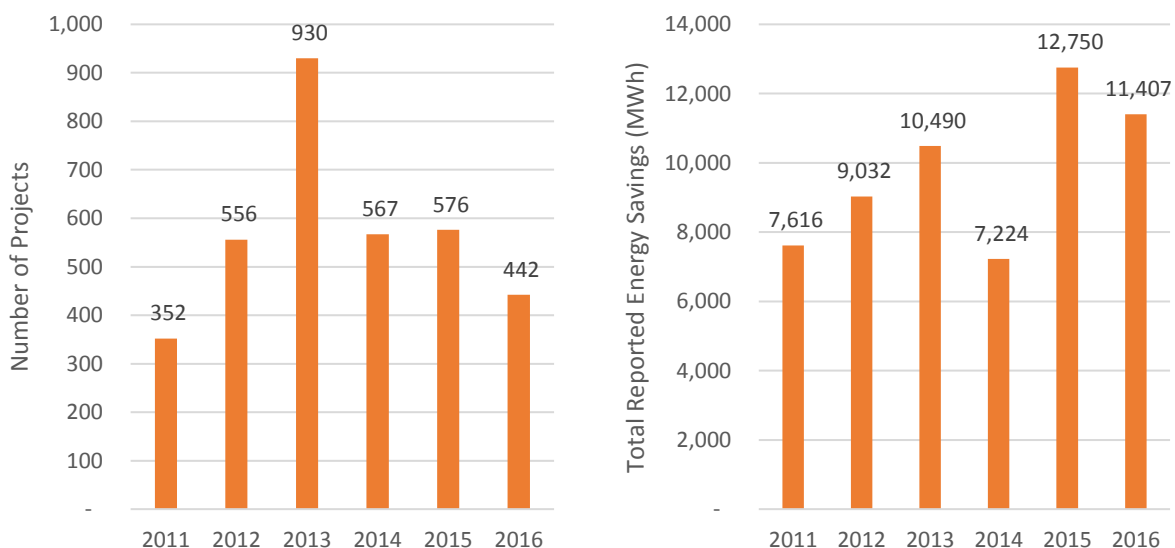

Figure 1. Participation summary from 2011 through 2016 with project count and reported savings.

The two major objectives of the evaluation were to (1) quantify energy savings on an annual basis and (2) determine drivers of program realization rates and identify ways the program can be improved. The authors considered several IPMVP[1] options to evaluate the program, and ultimately decided to follow IPMVP Option C and utilized customer billing data to develop a robust regression model to evaluate pre-/post- energy consumption. To assess the drivers of the energy realization rate, additional primary and secondary research was performed.

The purpose of this paper is to take a holistic look at key program implementation changes going back to 2011 and highlight the changes that led to increased program realization rates from 39 percent in 2011 to 58 percent in 2013 to 80 percent in 2016. An econometric analysis that uses participants as their own controls for savings estimates by employing a variation in adoption model was the basis for verified energy savings and program realization rates. Using participants and in-the-pipeline participants as a comparison group incorporates some aspects of program attribution and other exogenous effects that cannot be captured from an engineering-based analysis alone.

Evaluation Methodology

The authors have evaluated a Small Business-focused program via econometric modelling of customer billing data annually since 2011. These findings were supported by on-site data collection, which was used both qualitatively and quantitatively to inform the verified savings from the econometric model. These findings shed light on and seek to provide insight on the engineering-based assumptions used to develop the utility reported savings estimates.

Billing analysis is common and considered as a favorable method for evaluating small business programs in National Renewable Energy Laboratory's (NREL) Uniform Methods Project, which is a credible methods guide for determining energy efficiency savings for specific measures. Variation in adoption models, which are a type of billing analysis, are one of the most commonly used methods in estimating energy efficiency impacts. This modelling framework also minimizes the self-selection bias compared to other billing analysis methodologies such as "matching".

The basis for the program *ex ante* reported savings were driven by 1) estimated connected load of existing fixtures; 2) prescriptive power demand estimates for new equipment; 3) self-reported hours of use; and 4) secondary effects from HVAC equipment. These engineering-based estimates did not account for additional factors such as operational or behavioral changes associated with efficient equipment, variations in baseline power demand, and other factors that the billing analysis inherently captures.

The diagram below highlights the data sources and evaluation processes used by the authors to determine ex post verified energy savings. Further detail on each data source and research activity are outlined below in Figure 2.

---

[1] The IPMVP (International Performance Measurement & Verification Protocol) outline industry guidelines and best practices for evaluations. Available at http://www.nrel.gov/docs/fy02osti/31505.pdf
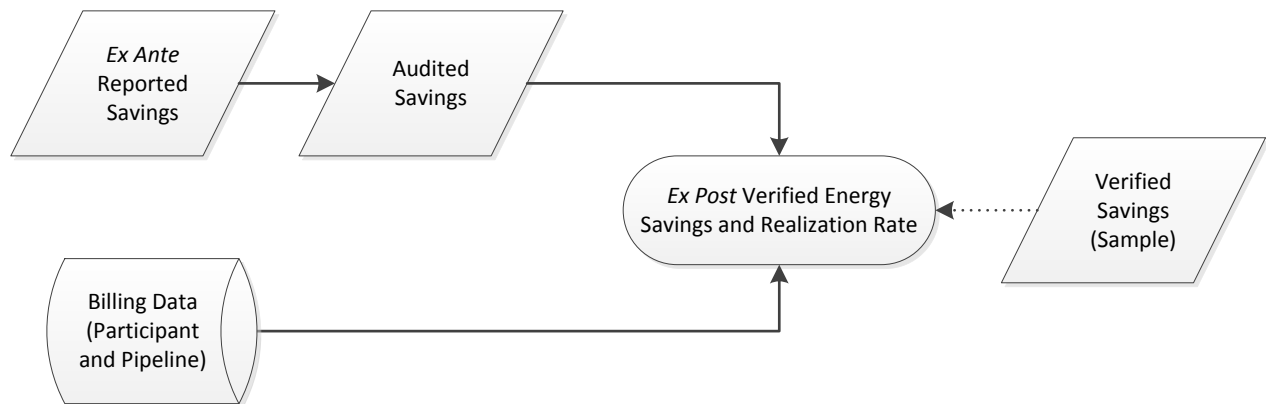
Figure 2. Overview of data sources and evaluation process.

Billing Analysis

A billing analysis of participants and pipeline participants serves as the basis for determining program energy savings. The authors employed a Variation in Adoption (VIA) method to estimate program savings. The VIA model takes advantage of the differential timing of program enrollment to develop a counterfactual and to identify program savings. The model essentially takes the perspective that the best comparison group for participants consists of those customers that enroll in the program in a later period, as shown below in Figure 3. These customers serve as the most suitable comparison group customers because these customers represent the counterfactual. In other words, future participants are assumed to be very similar to the program participants prior to the program intervention. This is also consistent with the average monthly consumption data, which indicates usage for participants and future participants are very similar prior to participation in the program.
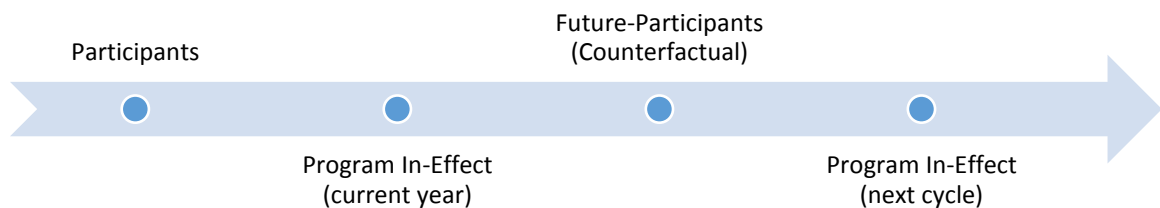

Figure 3. Illustration of billing analysis participation timeline.

The use of pipeline participants as a comparison group accounts for other exogenous effects, such as macro-economic trends. Pre- and post-installation periods are determined on a project-by-project basis. This modelling framework also takes project specific characteristics into account. These characteristics are non-time related but usually unobserved with billing data such as square footage of the premise. Because the billing analysis does not consider time of day savings, the demand savings are not verified via billing analysis.

The regression analysis accounts for seasonality of savings due to HVAC interaction effects via the inclusion of seasonal binary variables. Program savings are estimated using a Variation-in-Adoption model, which relies only on program participants to develop the counterfactual. Customers who participate in the program later serve as the control group for customers who participate in the program early on. This model also accounts for the variation in project size in the regression equation.

A study on the Business Direct Install Program in Massachusetts employed an econometric model to estimate the program savings and the associated realization rate. The overall estimated

realization rate was 66 percent, although the realization rates varied by measure category. For example, the realization rate for lighting fixtures were 67 percent while the realization rate for refrigeration was 89 percent[2]. The study authors also noted that they believed this analysis likely underestimated program savings, and that a similar billing analysis between 2003 and 2007 found realization rates ranging from 86 percent to 104 percent. Without detailed regression outputs, it is difficult to understand the issues faced in these studies; the authors would simply note that further research is required to understand specific differences in modelling techniques and drivers of the realization rates.

Another study evaluating the energy savings for the Better Buildings Neighborhood Program using billing analysis employed a sector-level regression model. Their findings showed 66 percent realization rate for commercial buildings[3]. In isolation, individual billing analysis provide meaningful evaluated energy savings estimates, but inherently lack context on the drivers on realization rates. The model relies on the assumption that, controlling for both customer and monthly fixed effects, neither energy use in month t, nor energy savings s months into the program, are correlated with the timing of program entry. Formally, the regression equation is given by:

Equation 1. Regression analysis: variation in adoption model.

$$ADU_{it} = \alpha_i + \beta_s * SeasonYear_{st} + \sum_j \gamma_{j,s} * PostSeasonYear_{j,st} + \varepsilon_{it}$$

Where i indicates the premise, t indicates the bill period, s indicates the season-year, j indicates the season, and

ADU_{it}	= Average daily usage (kWh) for premise i in period t
α_i	= The constant term ("fixed effect") for premise i
$SeasonYear_{st}$	= A series of binary variables taking a value of 1 if period t is in season-years. Fall is considered the baseline season because it is the first complete season of the analysis period.
$PostSeasonYear_{it}$	= A binary variable taking a value of 1 if the measure has been installed at premise i prior to period t for each season during each evaluation year.
ε_{it}	= The model error for participant i in period t. Standard errors are clustered to account for heteroskedasticity and autocorrelation at the participant level.
β_s, γ_s	= Model parameters

Seasons are defined by the following cut-off dates:
Winter   January 1 – March 31
Spring   April 1 – June 30
Summer   July 1 – September 30
Fall   October 1 – December 31

The authors expect slight seasonal variation of savings due to the interaction effects between lighting and the HVAC system. To account for the seasonality of savings, the authors included seasonal

---

[2] This study produced measure-type specific realization rates, unlike the model used in this study. Available at: http://ma-eeac.org/wordpress/wp-content/uploads/Massachusetts-Small-Business-Direct-Install_2010-2012-Impact-Evaluations-1.29.13.pdf
[3] This realization rate is a blended value from site verification and billing analysis realization rates, with the actual billing analysis realization rate at 21%. Available at: https://www1.eere.energy.gov/analysis/pdfs/bbnp_volume_1_final_evaluation_072215.pdf

binary variables. Seasonal binary variables allow energy usage to vary by season without impacting the overall savings estimate. These variables are sufficiently flexible to capture the effects of changes in weather and other factors that change by season, such as extended business hours during a holiday season.

Annual savings for the program each year are calculated as the estimated average energy use during the post-installation periods multiplied by the project count. This was compared to the ex ante savings estimates in order to calculate a program realization rate.

Parameter estimates for variables involving post participation energy use are negative, indicating usage decreases after program measures have been installed. T-statistics greater than 1.64 indicate the parameter is statistically significantly different from zero at the 90 percent confidence level. In particular, the team ensured that the post savings parameter, representing the estimated realization rate, is statistically significant each year.

Alternative Specifications of the Model Over Time

Each year the team worked through several iterations of the model in an effort to reduce uncertainty while ensuring that any additional controls or data manipulations are logical and consistent with participation characteristics. For example, the team explored the following model adjustments:

- Binning participants based on either project savings or total facility usage, and running separate models for each bin.
- Exploring statistically adjusted engineering (SAE) models, which incorporate the ex ante claimed savings for each project in the regression.
- Exploring changes to season definitions.
- Variables to control correlation between participation date and facility size (e.g., larger facilities reported in Q4 vs. Q1.)
- Data normalization techniques to model either total kWh or realization rate directly.
- Requests for additional billing data depending on Q4 p-values.
- Removal of sites considered to be outliers.

The authors find that these additional model adjustments generally had minor (less than 5 percent) effects on the realization rates and confidence intervals. For consistency, the team opted to utilize the model specification outlined above, subject to minor adjustments[4], each year.

Engineering Review and On-site Verification

Because the billing analysis is not able to discern the drivers of the program realization rates, the authors performed additional research to substantiate and explain the billing analysis results. This research included both engineering reviews of all assumptions, and limited on-site verification visits. The energy savings algorithm used to calculate engineering-based savings estimate is shown below; all parameters were explored either through secondary research, and/or primary on-site visits.

Equation 2. Ex ante energy savings algorithm.

$$kWh\_savings = [kW\_base * (QTY\_base - QTY\_burnout) - kW\_eff * QTY\_eff] * Hours * HVAC\_kWh$$

---

[4] Minor adjustments may include including additional monthly dummies or interaction terms. These variables are included to improve the robustness of the model and should not change the program energy impacts.

Where:

kWh _savings = energy savings (kWh)
kW _base = connected load (kW) of baseline equipment
kW _eff = connected load (kW) of efficient equipment
QTY _base = quantity of baseline fixtures
QTY _burnout = quantity of non-operational baseline fixtures
QTY _eff = quantity of efficient fixtures
Hours = estimated annual hours of use
HVAC _kWh = energy interactive effect

The authors compared the reported parameters with the parameters used in other business program evaluations in Ohio and in other jurisdictions, and findings from the on-site verification work. The on-site verification site visits, approximately 20 each year, were designed to both gather primary information on program operation and solicit feedback from customers, as well as verify quantities, locations, wattages, and space types associated with each measure. The authors also solicited information related to operating schedules to assess the quality of the hours of use estimates provided in the tracking data. These visits were primarily performed to gather both qualitative and quantitative data to help understand the drivers of the realization rates from the billing analysis. The team used engineering judgement to develop recommendations based on the on-site findings.

Challenges and How They Were Addressed

Every impact evaluation has strengths and weaknesses; the evaluator's goal is to minimize the weaknesses inherent with a particular methodology while taking advantage of the strengths. The billing analysis described in this paper has the benefit of working with the entire population of participant data, but one downside is that the findings are aggregated at the program level. This can make it difficult to understand drivers of program performance and identify shortcomings of the ex ante engineering-based assumptions.

Demand Savings

The data available to the evaluation team in Ohio consisted of monthly billing data, which precludes the possibility of estimating robust demand savings during the utility peak period. Other jurisdictions with AMI meters and access to hourly data should allow for estimating program-wide demand savings.

Uncertainty

The authors quantified uncertainty in the econometric model by calculating a relative precision at the 90 percent confidence level. This confidence interval is fundamentally different from the sampling error often reported for impact analyses that employ a sampling approach to develop verified savings estimates, but nevertheless is useful to assessing the predictive power of the model.

Other Effects on Energy Consumption

Because the billing analysis is informed by the total facility energy consumption, there are often other factors that may affect energy consumption beyond those directly related to participation in the energy efficiency programs. These can include:

- Behavioral changes, such as snap back, changes to how lighting is used and controlled, thermostat set points, etc.
- Facility operational load changes, including increased production, spillover effects, changes to plug loads and other equipment, etc.
- Weather events.
- Seasonality of operation.

By utilizing the VIA model rather than a matched-control group model, the authors believe that these effects are largely reduced, although not entirely removed. Because these effects can produce either increased or decreased load, it is reasonable to assume that at a minimum some of these effects cancel out. As opposed to matched-control group model, the VIA model does not use a pool of non-participants to develop a comparison group. Conversely, the VIA model uses future program participants to serve as a control customers. The matched comparison group approach tries to develop a comparison group using non-participants and matches participants to non-participants based on their usage during the pre-program period. The robustness of the program impacts heavily relies on the quality of the matched comparison group. However, the VIA model uses future participants as the control, who have very similar characteristics but have not been exposed to the program yet. This reduces the selection-bias significantly.

Findings

The results revealed that the initial ex ante engineering assumptions were significantly overstating energy savings estimates. Reducing these parameters with the application of realization rates resulted in savings estimates more in line with those supported by the billing analysis. By continuing to explore the assumptions over time, the program benefitted from reduced uncertainty around program savings by having a statistically valid econometric analysis informing overall performance.

The key program improvements implemented since 2011 to increase realization rates include the following, which are outlined in more detail below.

- Implementing more reasonable hours of use estimates specific to each fixture.
- Capturing the role of lamp burnouts.
- Reporting of HVAC interactive effects in unconditioned or exterior spaces.
- Reviewing and updating baseline and efficient fixture wattages.
- Incorporating safeguards when project savings exceed reasonable limits.
- Ensuring that change orders from installations are captured in program tracking databases.

The authors have tracked key program changes since 2011, along with findings and recommendations made throughout the years.

Realization Rates from 2011-2016

The realization rates for each year from 2011 through 2016 are shown in Figure 4 below, along with a visual representation of the 90 percent confidence interval. Overall, realization rates are trending upwards as the program continues to mature and make targeted improvements in the accuracy of reported energy savings. The authors identified several key drivers of the realization rates, which are explored below.
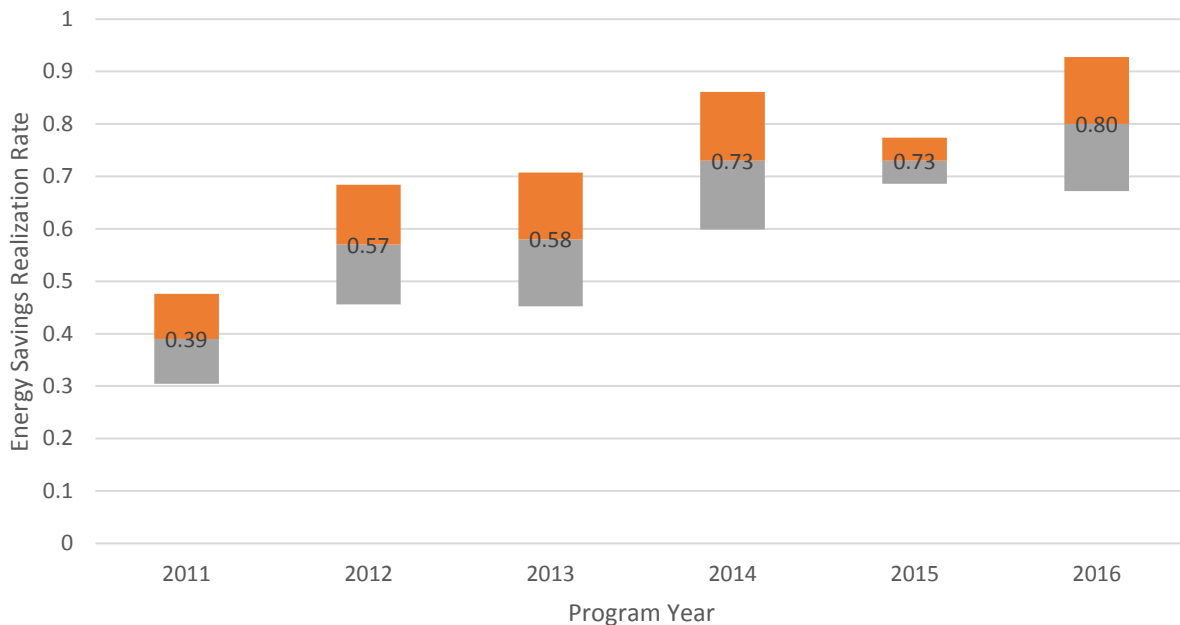
Figure 4. Energy savings realization rate by program year, with 90 percent confidence interval.

Hours of Use

The estimated hours of use variable was identified in 2011 as a key driver of the realization rate. The form used to report hours of use for a facility was not granular enough to provide meaningful hours of use estimates, which resulted in apparent overestimation of hours of use. The recording method for hours of use was overhauled as part of the program relaunch in 2013, where fixture-specific hours of use were determined by the auditor with input from the customer on operational characteristics. From 2014 through 2016, the implementation contractor continued auditor training to ensure consistency and accuracy of estimates between individual auditors. The average hours of use have decreased over the years and is closer to established deemed building hours of use.

HVAC Interactive Effects

HVAC interactive effects reflect the reduced load on the facility's cooling system from efficient lighting. This can increase the energy savings associated with efficient lighting in conditioned spaces beyond the energy savings accrued by the equipment itself. The program initially reported a single deemed facility-wide HVAC interactive effect for all fixtures installed, regardless of whether the space was conditioned. For unconditioned or exterior spaces, this results in an overestimation of savings.

Starting in 2013, as the program began to track installation location in the new auditing tool, the team now had access to significantly better information that could be used to assess whether HVAC interactive effects were applied correctly. The authors found that in many cases individual fixtures were being credited with HVAC interactive effects mistakenly, for example in spaces labeled "Exterior" or "Outside". As the program began to roll out more LED lighting measures in 2014 and 2015, this became a larger issue because more lighting retrofits occurred in high bay, unconditioned, and exterior locations. The IC has implemented additional auditor training and refined their tools in 2015 and 2016 to better assign HVAC interactive effects.

As to the HVAC interactive effects themselves, there is uncertainty around these values that stem from the modelling practices used to generate savings. Specifically, the values are derived from primarily large stock building types from DOE-2 or EnergyPlus simulations, and may not be accurate for the small business segment. Most small business fall under the "Small Retail" or "Small Service" designation, which may not accurately represent the corner grocers, gas stations, convenience stores, small offices, small manufacturing and ware facilities, and other miscellaneous building stocks prevalent in the small business space.

Lamp Burnouts

Existing lamp burnouts were identified as a potential issue in 2011. The billing analysis implicitly captures reduced energy consumption from burnouts in the pre-retrofit condition, which can result in significantly less savings on the customer bill post-retrofit. The authors found that small businesses often have a significant tolerance for burnouts, likely driven by the lack of dedicated maintenance staff. By tracking burnouts in the initial audit, the program can ensure that the estimated savings will more closely mirror actual savings experienced by the customer. This has a side benefit of better managing customer expectations. In 2013 the IC began to track burnouts found in the initial audit, and implemented training for auditors to ensure that they are capturing this data correctly.

Wattage Assumptions

The authors could verify equipment wattages based on manufacturer specifications during the limited on-site data collection effort each year. From 2011 through 2014, most fixtures installed through the program were T8 linear fluorescent lamps and electronic ballasts, and from 2015 onward most fixtures were LED, either linear systems with electronic ballasts or integrated exterior/high-bay fixtures.

The authors found through onsite visits in 2013 and 2014 that often 32W T8 lamps were installed where 28W lamps were requested by the auditor and found in the tracking data. The IC increased installation contractor training to ensure that correct lamps were installed; this issue was less common in 2014, and in 2015 the majority of linear systems had shifted to LEDs.

There is additional uncertainty surrounding the connected load for linear LED system with separate ballasts, as not every configuration is explicitly detailed in manufacturer specifications. The authors recommended working to ensure that total installed fixtures wattages were accurate.

The baseline wattage, on the other hand, is impossible for the evaluation team to assess. The authors in 2013 compared lamp and ballast wattages to deemed values from the Ohio Technical Reference Manual (TRM). As shown below in Figure 5, for the most common 4-foot and 8-foot T8 retrofits, reported baseline wattages were generally higher than the TRM values. The IC has fine-tuned their audit practices and baseline assumptions to reflect the lower wattages typically found on site.
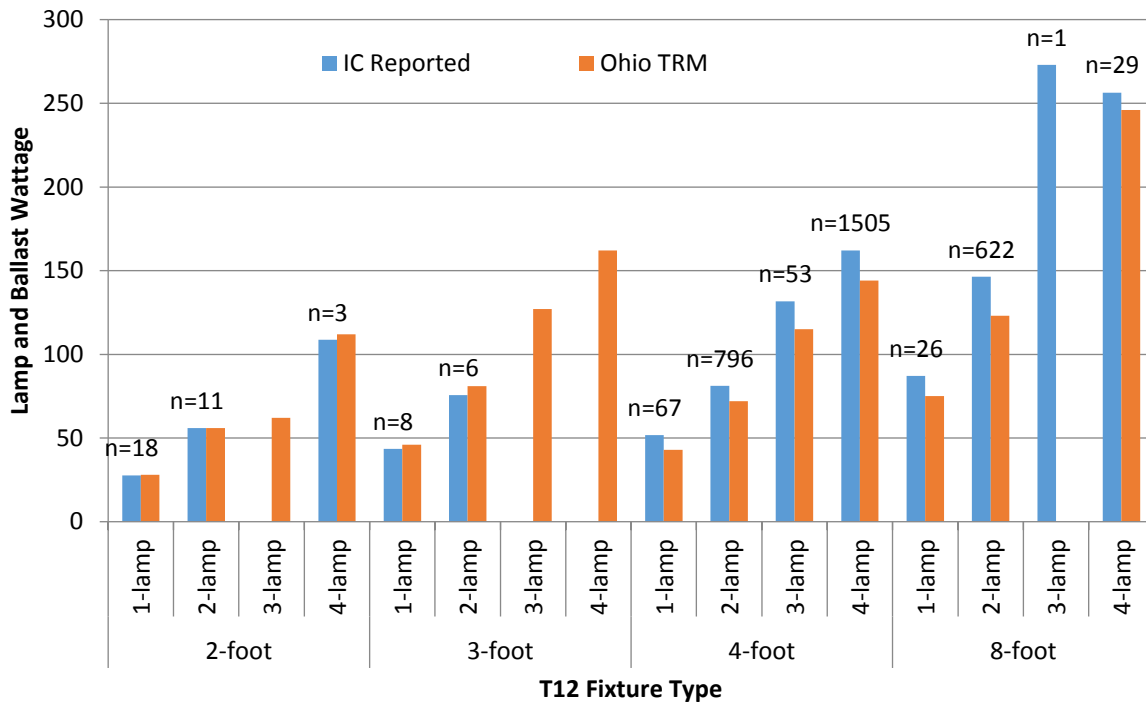
Figure 5. Comparison of 2013 baseline wattages.

Quantity

From on-site verification work between 2011 and 2016, installation rates and equipment quantities were not identified as a meaningful driver of the realization rate.

Conclusions

AEP Ohio's Small Business Program demonstrates that incremental improvements can be achieved and measured over time through an annual billing analysis and supplementary primary and secondary research. The gross realization rate estimated through the billing analysis serves as a barometer of program performance and a high-level check of deemed savings assumptions, while engineering judgment and an understanding of the program implementation characteristics drive targeted recommendations for improvement. The authors intend to continue improving the program using this methodology, and identifying areas for improvement and targeted research to both explain realization rate drivers and increase program realization rates above 80 percent.

The econometric model used for this billing analysis is a VIA model, which the authors believe best estimate gross energy savings for the population of program participants by utilizing participants as their own control group.

The key improvements and recommendations made from 2011 through 2016 include 1) introducing a method for accounting for lamp burnouts; 2) implementing more reasonable hours of use estimates specific to each fixture; 3) removing HVAC interactive effects from fixtures in unconditioned spaces; 4) revisiting baseline and efficient wattage assumptions; 5) incorporating safeguards when project savings exceeds reasonable limits; and 6) ensuring that change orders from installations are captured in program tracking databases. These improvements have increased realization rates from 0.39 in 2011 to 0.80 in 2016.

Using a consistent billing analysis method year over year, supplemented with limited onsite work and supplementary research, allows for a cost-effective and meaningful small business program evaluation. Continuity of evaluation approach and continual improvement from program managers and implementers allow for longitudinal comparisons that result in more effective energy efficiency programs.

References

Efficiency Valuation Organization (EVO). International Performance Measurement & Verification Protocol. Concepts and Options for Determining Energy and Water Savings Volume I. Available at: http://www.nrel.gov/docs/fy02osti/31505.pdf

Cadmus 2013. Massachusetts Small Business Direct Install: 2010-2012 Impact Evaluations. Available at: http://ma-eeac.org/wordpress/wp-content/uploads/Massachusetts-Small-Business-Direct-Install_2010-2012-Impact-Evaluations-1.29.13.pdf

United States Department of Energy of Efficiency and Renewable Energy, 2015. Evaluation of the Better Buildings Neighborhood Program: Final Synthesis Report, Volume 1. American Recovery and Reinvestment Act of 2009. Available at: https://www1.eere.energy.gov/analysis/pdfs/bbnp_volume_1_final_evaluation_072215.pdf

Department of Energy 2013. The Uniform methods Project: methods for Determining Energy Efficiency Savings for Specific Measures. Available at: https://energy.gov/oe/downloads/uniform-methods-project-methods-determining-energy-efficiency-savings-specific-measures