Are Savings from Behavior Programs Ready for TRM Prime Time?

Katie Parkinson and Scott Dimetrosky, Apex Analytics, Boulder, CO
Jeffrey Orrey, GeoVisual Analytics, Boulder, CO

ABSTRACT

Behavior-based programs are increasing in number and magnitude throughout the country. Concurrent with this trend, more and more regions are developing Technical Reference Manuals (TRMs) as a way to provide more consistent savings assumptions across different program administrators. The paper summarizes savings claims from information "opt-out" behavior-based programs across the country, finding that the savings range from a low of 1.2% to a high of 3.3%. Given the range of savings, the paper does not recommend that TRMs deem savings from behavior-based programs. However, there are different methods that may be employed to determine the savings, and these methods can impact the savings claim. Furthermore, there are a number of policy decisions around the savings claims, including measure lifetime (persistence), double counting, monthly vs. annual claims, and measure incremental cost. Given the uncertainty around the methods and the policy decisions, the paper recommends that TRMs adopt a methods protocol for behavior-based programs. The paper also recommends that more research be conducted around how the different methods may drive savings upwards or downwards.

Introduction

Behavior-based programs are increasing in number and magnitude throughout the country. OPower, the largest provider of behavior-based programs, now serves 85 utilities in 30 US states and seven countries and has grown by almost 200% over the past three years.[1]

Concurrent with the growth of behavior-based programs, many regions of the country are developing Technical Reference Manuals (TRMs). As of April 2013, there were approximately17 states with energy efficiency programs that were impacted by TRMs. TRMs serve a number of purposes, including:[2]

- Provide consistency and transparency, across multiple program administrators, regarding savings assumptions and calculations and the underlying sources of those assumptions and calculations.
- Support the calculation of the benefit cost tests in support of program design, evaluation and regulatory compliance.
- Identify gaps in robust, primary data, that can be addressed via evaluation efforts and/or other targeted end-use studies.
- Provide a process for periodically updating and maintaining records, and preserve a clear record of what deemed parameters are/were in effect at what times to facilitate evaluation and data accuracy reviews.
- In some regions, to provide certainty as to savings values that will be used in assessing performance against energy efficiency or capacity savings goals.

Since not every measure can be included in a TRM, there is usually a process for selecting measures that represent either significant claimed savings or large potential savings. Behavioral-based

[1] Provided by OPower in email communication, April 2013.
[2] These objectives are modified from the 2012 Illinois TRM.

programs, as one of the fastest growing opportunities for energy savings, are beginning to meet both of these criteria in many regions in the U.S. A review of the current TRMs, however, revealed only one that included actual assumptions for savings values (Massachusetts) and two that included evaluation protocols (Arkansas and Pennsylvania).

This paper explores the question as to whether or not more TRMs should include behavior-based programs. First, the paper explores whether or not there is consistency in energy savings across regions; next, the paper delves into some of the methodological questions regarding estimating savings from behavior-based programs; finally, the paper makes a recommendation as to how behavior-based programs should be treated in TRMs.

Variation in Savings Estimates

To explore the range of savings estimates for behavior-based programs, the authors collected recent evaluations of 27 residential electric programs that had deployed OPower during the last four years.[3] This study limited the review to OPower electric programs for a number of reasons, including:

- OPower has, to date, been the largest provider of behavior-based programs;
- There were far more available studies for electric programs than for gas programs;
- Variability could be minimized by limiting the analysis to one provider, one program type (an "opt out" approach where energy benchmarking information was included on customer bills), and one fuel type. If the variability proved acceptable then the analysis could then be expanded outward by adding in additional providers and gas savings.

As shown in Figure 1, the Average Treatment Effects (ATE) ranged from a low of 1.2% to a high of 3.3%. There was no real clear pattern in savings by year nor by region of deployment. Table 1 summarizes the findings. On average, OPower electric deployments reduced energy bills by 2.1%, with a relative precision (at the 90% confidence level) of this estimate of ± 15%. This means that, with 90% confidence, the savings could be as low as 1.9% or as high as 2.3%.

While these may appear to be small differences, these percentages are being calculated based on total energy usage, and thus can differ in absolute magnitude quite significantly. For perspective, many utilities around the country have ***total*** energy savings targets of 1% to 1.5% of their sales. A difference of just a tenth of a percent, when extrapolated to tens of thousands of customers, can represent quite a large difference in energy savings.

[3] Note that 14 of the studies are summarized in a 2011 Journal of Public Economics article by Hunt Allcott. The remaining studies were collected and analyzed by Apex Analytics. Citations for these studies are provided in the Bibliography.

Figure 1. Average Treatment Effects for OPower Electric Deployments, 2009-2011

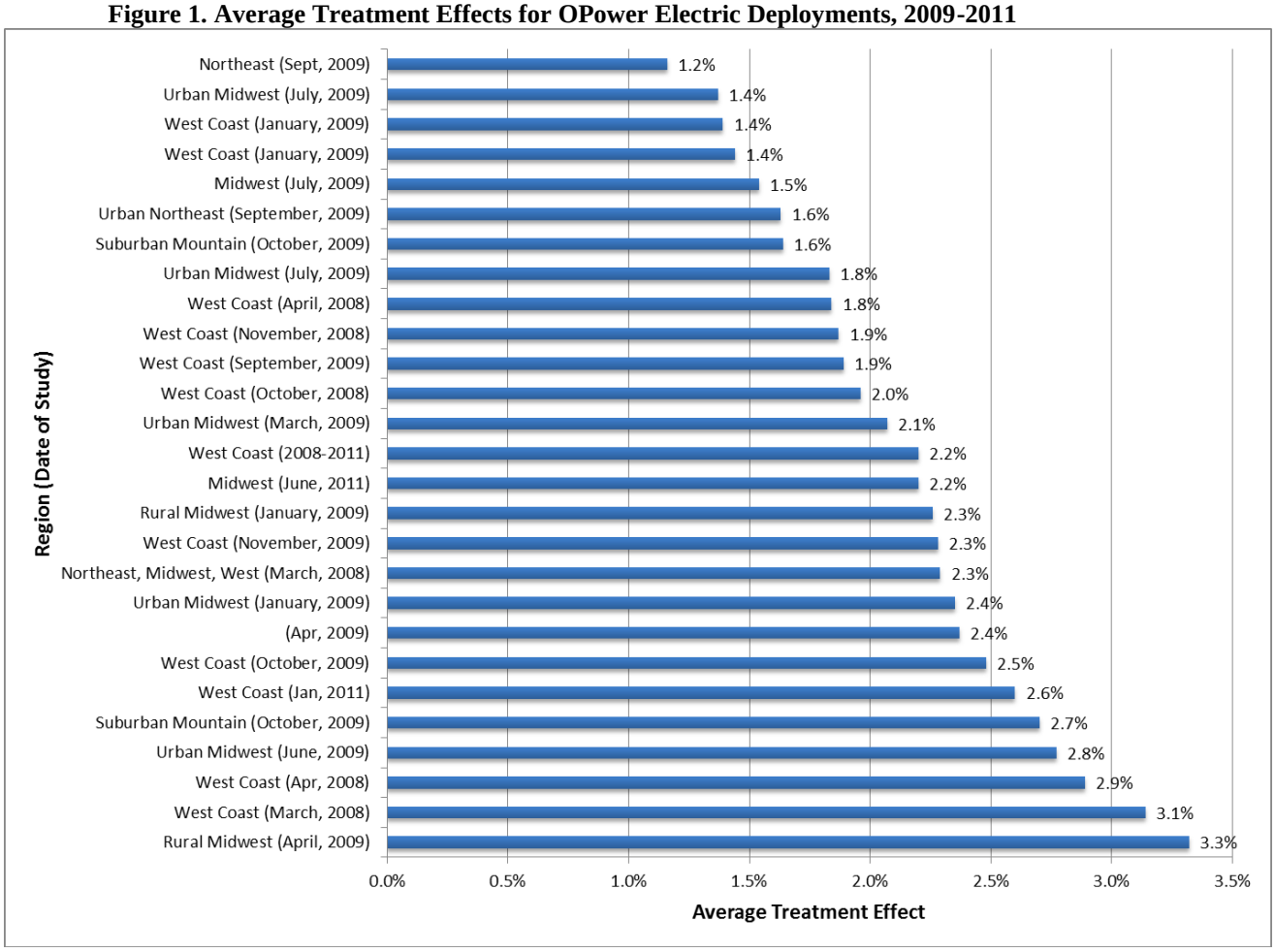

Table 1. Summary of Estimated Savings

		Confidence Intervals (90% CI Level)			
Number of Studies	**Savings Range**	**Lower Bound**	**Average Savings**	**Upper Bound**	**Relative Precision**
27	1.1%-3.3%	2.0%	2.1%	2.3%	15.1%

Variation in Methodological Approaches

While all the studies cited above utilize a billing analysis to determine energy savings, there are a number of methodological decisions that can vary between studies. Note the purpose of presenting these is not to recommend one approach over another, but to understand the potential differences in approaches that may lead to potentially different savings outcomes. A recent publication from the State and Local Energy Efficiency (SEE) Action Network discusses each of these methodological points in detail (LBNL, 2012).

Study Design

Evaluations must rely on either experimental or quasi-experimental design to determine savings. In experimental design, customers will be placed in either a treatment or control group based on a randomized control trial (RCT) approach. Quasi-experimental designs include methods such as the regression discontinuity method (selecting a comparison group based on customers that meet similar eligibility requirements) and the variation in adoption approach (selecting participants that enroll later in a program as a control group for those that participated in a prior period), but both of these introduce additional error into the analysis.

Length of Study and Baseline Period

Reporting deadlines and data availability may limit the length of the study or baseline periods. For example, many annual energy efficiency reports are due in the first or second quarter of the year, which may put pressure on evaluators to use a truncated study period. In other cases, pulling billing system data from more than a year or two back may require accessing archive data, which may be cost-prohibitive.

Annualizing Savings

Most jurisdictions in the U.S. report first year annualized savings, meaning that full year annualized savings are reported even for measures that get installed for only part of the year. Evaluations of behavior-program evaluations, however, often report monthly energy savings (since the models are typically based on monthly billing data). A utility annualizing savings from behavior-based programs could report significantly greater savings than a utility that limits the reporting to participant months (i.e., the number of participants for each month that received energy reports).

Panel vs. Aggregated Data

Analysis models can either use energy data that are aggregated across time for both the pre- and post-program periods (e.g., average energy use for the period prior to and during the program year) or panel data (also called time series of cross-sectional data), which typically are data from multiple time points for pre- and post-program periods (e.g., monthly energy use for each month of the program and for the twelve months preceding the program).

Comparison of Energy Usage vs. Change in Energy Usage

Analysis models can be specified to estimate program savings by either comparing the energy saved by the treatment group to the energy saved by the control group (a "difference in differences" approach), or comparing the energy use of the treatment group to the control group during the program (a post-treatment comparison).

Equivalency Check

The degree to which a savings estimate is unbiased depends on how similar the control group is to the treatment group. Some studies determine if the households in the treatment group have characteristics that are statistically similar to those in the control group (e.g., compare the monthly or yearly pre-program energy use and the distribution of pre-program energy use between the treatment and control group), while others do not.

Treatment of Households that Opt Out or Close Accounts

The treatment of households that opt out may vary between studies, with some studies leaving these homes in and others taking them out. This could potentially bias the results (e.g., households that drop out may take less actions to reduce energy than those that continue to receive the home energy

reports). The treatment of households that close their accounts – including any changes in tenants at the same site – may also vary between studies.

Controlling for Double Counting

For programs in which efficiency measures can be tracked to a specific household (e.g., installation of insulation by a contractor), double-counted savings can be directly determined as the incremental participation in the non-behavior program. While most program evaluations to date have assigned all the incremental double counted savings to the other (non-behavioral) program, there is increasing recognition and acceptance that the behavior programs helped cause these additional savings, and thus could be credited with some percentage of these savings. Differences in the handling of double counted savings could lead to significantly different savings estimates for the same program.

Expected Useful Life/Persistence

There is substantial debate and ongoing research regarding the Expected Useful Life (EUL) and persistence of energy savings from behavior-based programs. There are two key aspects in this discussion. The first aspect is how savings will vary from when participants first receive reports to when they receive reports after many months or even over a year or two. The second aspect is how long, if at all, savings might continue after participants stop receiving home energy reports.

Summary of Findings

This paper examined both variation in savings and variations in methods for behavior-based programs. The examination of savings found that, while there is a wide range of savings (from 1.2% to 3.3%) the average saving was 2.1%, with moderate precision levels (± 15%). Even small differences in the percent saving, however, can translate to significant savings when multiplied by tens of thousands of customers.

In terms of variation in methods, there are a number of methodological decisions that will lead to significantly different savings assumptions. These decisions range from the study design (e.g., experimental vs. quasi-experimental design and length of study period), to data analysis decisions (e.g., comparison of energy usage vs. change in energy usage, handling of opt out households) to policy decisions (e.g., controlling for double counting).

Conclusion: Should Behavior Program Savings be Included in TRMs?

As noted above, one of the primary goals of a TRM is to provide transparency and consistency in the savings estimates. This paper demonstrates that there is a fairly wide range of savings estimates for behavior-based programs, plus there are many methodological decisions that can drive the ultimate outcome for measure savings. In fact, it is quite possible that one reason for the range of savings estimates is that the methodology adopted between many of these studies differed. Unfortunately, a number of reports did not contain sufficient detail to determine the exact details on the methodology to determine if this assumption of causation is valid, but the findings are suggestive that the differences in methodology at least contributed to the differences in savings estimates. The authors of this paper highly recommend that more research be conducted to understand exactly how the differences in methodology lead to differences in expected savings.

Based on these findings, it appears that, at a minimum, methods protocols for behavior-based programs should be included in TRMs. This will accomplish a number of objectives, including:

- Ensuring that recommended methodologies are utilized, and that sub-optimal techniques and shortcuts (e.g., shorter analytical periods) are avoided;

- Providing for more consistent approaches between program evaluators, which will allow for improved validity in comparing findings from one study to another;
- Allowing for more transparency and understanding of how the results were derived, particularly if reporting requirements are also included as part of the protocol development.

The authors also believe that it is likely that, in the near future, the savings values may be stable enough that they too can be deemed in TRMs. The first step in this direction might be deeming savings values based on specific program design and implementation, basing the savings off of prior research findings.[4] This approach also accounts for potential regional differences (e.g., weather impacts) or implementation differences (e.g., targeting high vs. moderate energy user). As methodologies become more consistent and savings across regions become more comparable it then may be possible disaggregate legitimate regional/implementation differences from methodological differences, and to deem prospective savings estimates for similar program deployment (e.g., identical home energy reports and targeted customer types) in different regions.

Will Deeming Methodology Stifle Research Innovation?

One concern with including behavioral-based programs in a TRM is that it will stifle research innovation. This same argument, in fact, is used against TRMs in general: once savings/methods are deemed, will it lead to less research (or less innovative research) to figure out what these measures/programs are actually savings? This argument is especially poignant with behavior-based programs, however, as they are still perceived as new programs with many areas of uncertainty (e.g., the breakdown of savings from behaviors vs. measure installations).

Including methodological protocols in TRMs, however will reduce these areas of uncertainty by providing improved transparency and consistency: true differences in savings between regions, implementation contractors, and/or reporting strategies can be determined through vastly improved cross-study comparison. This is not only one of the primary purposes of a TRM, but also the goal of another recent Department of Energy (DOE) initiative called the Uniform Methods Project.[5]

In addition, TRMs need to be written in a way that they allow for improved methods/approaches, with the understanding that these changes may need to get applied prospectively, or even retrospectively, depending on the policy of the specific TRM. This is no different from a TRM algorithm that might be used for an air-conditioner or CFLs: as improved information becomes available (e.g., metered data on effective full load hours or run times) these data get incorporated into the next iteration of the TRM. TRMs are living documents, and can incorporate improved models/methodologies as research demonstrates they are improvements over existing approaches, just like they incorporate changes to other parameter inputs.

Behavior-based programs, therefore, are ready for TRM prime time in the form of methodological protocols. Once this step is accomplished, it could pave the way for behavior-based programs to be included in TRMs with deemed savings values.

[4] This approach was used in Massachusetts, which assumes savings of 1.38% to 2.1% based on the year and type of participant.

[5] More information on the Uniform Methods Project can be found at http://www1.eere.energy.gov/office_eere/de_ump.html

Bibliography

Allcott, H. 2011. "Social Norms and Energy Conservation." Journal of Public Economics.

Ayres I., S. Raseman, and A. Shih. 2009. "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage." *In Proceedings of the 5th Annual Conference on Empirical Legal Studies*.

Environmental Defense Fund. 2011. *Behavior and Energy Savings – Evidence from a Series of Experimental Interventions*.

Integral Analytics, Inc. 2012. *Impact & Persistence Evaluation Report to the Sacramento Municipal Utility District Home Energy Report Program*.

KEMA, Inc. 2012. *Puget Sound Energy's Home Energy Reports Program: Three Year Impact, Behavioral and Process Evaluation for Puget Sound Energy*.

KEMA, Inc. 2010. *Puget Sound Energy's Home Energy Reports Program: 20 Month Impact Evaluation for Puget Sound Energy*.

Todd, A., Stuart, E., Schiller, S., and Goldman, C. 2012. *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*. Berkeley, Calif. Lawrence Berkeley National Laboratory.

Navigant Consulting 2012. *Evaluation Report: Home Energy Reports. Energy Efficiency / Demand Response Plan: Plan Year 3 for Commonwealth Edison Company*.

Navigant Consulting 2012. *Evaluation Report: Home Energy Reports. Energy Efficiency / Demand Response Plan: Plan Year 4 for Commonwealth Edison Company*.

Navigant Consulting 2011. *Evaluation Report: OPower SMUD Pilot Year 2 for Sacramento Municipal Utility District*.

Navigant Consulting 2010. *Evaluation Report: OPower Pilot. Energy Efficiency / Demand Response Plan: Plan Year 2 for Commonwealth Edison Company*.

Opinion Dynamics Corporation and Navigant Consulting 2012. *Massachusetts Three-Year Cross-Cutting Behavioral Program Evaluation Integrated Report for the Massachusetts Energy Efficiency Advisory Council*.

Opinion Dynamics Corporation, Cadmus Group, Navigant Consulting and Michaels Engineering 2012. *Impact and Process Evaluation of 2011 (PY4) Ameren Illinois Company Behavioral Modification Program for Ameren Illinois Company*.

Power System Engineering, Inc. 2010. *Measurement and Verification Report of OPower Energy Efficiency Pilot Program for Connexus Energy*.

Summit Blue Consulting 2009. *Impact Evaluation of OPower SMUD Pilot Study for the Sacramento Municipal Utility District.*