

Causality, Sustainability and Scalability – What we still do and do not know about the impacts of comparative feedback programs

Mitchell Rosenberg, DNV KEMA Energy & Sustainability, Burlington, MA
G. Kennedy Agnew, DNV KEMA Energy & Sustainability, Madison, WI
Kathleen Gaffney, DNV KEMA Energy & Sustainability, Oakland, CA

Abstract

This paper gathers and organizes findings from publicly available large-scale, independent evaluations of comparative energy use feedback programs in the United States to assess how they have worked and to identify lessons learned for program design and evaluation. Comparative feedback programs such as Opower and C3 Residential provide monthly or quarterly reports to customers that compare their metered energy use to average consumption among their neighbors. These comparisons are meant to stimulate attention to energy use and adoption of energy efficiency measures and behaviors. The design of the reports relies on theories of behavioral influence that identify validation of recommended actions through reference to the actions of others in a similar situation—“social norming” -- as an effective strategy.

Assessment of evaluations reviewed here yields the following key conclusions:

- Customers assigned at random to receive comparative feedback reports reduce their annual consumption by a measurable amount, usually in the range of 1 – 3 percent, compared to counterparts who do not receive the reports.
- Savings persist and often grow through the second program year. In the third and fourth years, savings for a given participant cohort remain positive, but may decrease.
- The programs are cost effective under fairly conservative assumptions.
- Savings increase with frequency of reports.
- Programs have a modest impact on report recipients’ participation in other energy efficiency programs.
- Surveys of recipients and counterparts (control group members) do not yield large or consistent patterns of differences between them in energy-related behavior.
- Recipients find the comparative aspect of the reports useful, but evidence on customer perception of program influence is inconclusive.

Introduction

Comparative Energy Use Feedback Programs

Beginning in 2008, energy efficiency program sponsors in the United States began to field comparative energy use feedback programs, generally as one element of a more extensive portfolio of programs to promote the adoption of energy-efficient products, services, and behaviors among residential customers. These programs were designed to provide information and advice to customers, with the object of inducing them to adopt one or more activities out of a range of potential energy saving behaviors.

Comparative feedback programs are generally operated by turnkey program vendors. Opower is the most visible of these vendors and has had programs in the field the longest. Significant competition is developing from both start-up and established companies. The key elements of the program approach are fairly consistent from one vendor to another, and include the following.

- **Home energy reports with comparisons to neighbors’ energy use:** Periodically, the program vendor processes the client utility’s energy bills into individual home energy

reports for the customers in the program. The report contains the following features:

- A comparison of the customer's billed energy use in the prior month to average use by neighbors, defined operationally as customers within a certain distance of the customer's home.
 - A chart showing the same comparison for the previous twelve months.
 - A verbal or graphic description of the recipient's performance such as "Great, Good, Below Average" or a happy or sad emoticon.
 - Financial costs or benefits associated with actual consumption versus the neighbors' average.
 - Tips for reducing energy use.
- **On-going feedback.** Most programs provide customers with monthly reports. Program vendors and sponsors have experimented with different periods of reporting, including quarterly, to assess whether the frequency of the reports affects. They have also monitored consumption patterns among customers whose reports have been suspended for a longer period to assess the persistence of savings over time.
 - **Random assignment of customers to treatment groups.** The program design obviates the need for customers to make voluntary efforts to enroll, such as filling out an application, responding to a survey, or buying something. This feature enables random assignment of customers to a treatment group (those who receive the reports) and a control group (those who do not), and thus the application of experimental designs in assessing program effects. The expected levels of average energy savings from the program are small in comparison to baseline consumption levels. Therefore, the use of experimental designs to control for the effects of potential influences on consumption other than inclusion in the treatment group is critical for discerning program effects through analysis of billed consumption over time.

Utilities have responded strongly to the potential value offered by the comparative feedback report approach. Opower reports contracts with 50 utility companies in the United States and is currently initiating a pilot implementation with a company in the United Kingdom. C3 reports contracts with several major utilities in New York, Massachusetts, and California. These companies and others have begun to introduce new features and approaches to increase cost-effectiveness and other benefits for program sponsors. These changes include provisions for voluntary participation outside of initial trials, detailed on-line audits customizable to the customer's home, and feedback points redeemable for discounts on merchandise.

We note that comparative feedback reports are used in many energy efficiency programs that do not feature frequent feedback or random assignment of customers to treatment groups. These include so-called "opt-in" in which customers sign up for bill analysis services and advice on line or via other means and community programs of various kinds. (Ehrhardt-Martinez et al) All of this activity will produce a rich set of results regarding the response of various sets of customers defined by geography, climate, housing type, and baseline consumption patterns to a wide range of offerings. In the meantime, the available independent evaluations of feedback programs can help us understand what they have achieved to date, and what questions we should be asking of them in the future.

Findings of Completed Evaluations of Comparative Feedback Programs

To date, the sponsors of four large comparative feedback programs have released the results of professional, independent evaluations to the public. The sponsors of these programs are Sacramento Municipal Utilities Department (SMUD), Puget Sound Energy (PSE), Pacific Gas & Electric (PG&E) and a consortium of Massachusetts electric and gas utilities (MA). Each of these programs used the Opower platform; some sponsors in Massachusetts fielded programs developed by C3. In the paragraphs below, we summarize the methods and basic findings for these studies,

focusing on results that address the following questions of interest to energy efficiency program administrators.

- To what extent do reductions in energy use observed in the first year of participation persist in later years?
- What effect do changes in details of program deployment, such as the frequency and format of reports, have on savings achieved?
- Which customer attributes are associated with high levels of savings through participation in feedback programs? Can these differences be reflected in strategies to increase program savings and cost-effectiveness?
- Through what specific actions do program participants achieve energy savings?
- To what extent does information and feedback received through the program stimulate recipients to participate in other energy efficiency programs? Are these savings incremental to what the other programs would otherwise have achieved?

Study Methods

Table 1 displays information on the completed evaluations whose results are available to the public as of this writing.

Table 1. Overview of Feedback Program Studies

Sponsor/ (References)	Region	Evaluation Period	Fuels	Data Collection & Analysis	Sample Sizes
SMUD (Integral Analytics 2012)	Sacramento CA & environs	4/2008 – 9/2011	Electric only	Billing Analysis Treatment Control Customer Survey On-site Survey Cross participation records analysis	~35,000 ~50,000 800 30
Massachusetts (MA) (Opinion Dynamics Corporation & Navigant Consulting 2012)	Massachusetts	10/2009 – 12/2011	Electric & Gas	Billing Analysis Treatment Control Customer Survey Cross participation records analysis	~25,000 ~25,000 1,002
Puget Sound Energy (PSE) (DNV KEMA 2010; DNV KEMA 2012)	Pacific Northwest	10/2008 – 12/2011	Electric & Gas	Billing Analysis Treatment Control Customer Survey Cross participation records analysis	31,618 40,007 1,369
Pacific Gas & Electric Home Energy Reports Trials (Sullivan et al. 2013)	βwave: Central Valley, East Bay γwave: full service territory $\mathbf{1}$ wave: targeted usage strata	βwave: 8/2011 – 12/2012 γwave: 11/2011 – 12/2012 $\mathbf{1}$ wave: 2/2012 – 12/2012	Electric Only Electric & Gas Electric Only	Billing Analysis Treatment (γwave) Control (γwave) On-site survey (all waves) Cross participation analysis	205,000 205,000 701

The studies contained the following elements.

- **Analysis of bill data to estimate treatment effects.** All four studies used analysis of billing data to estimate savings associated with assignment to the treatment group. Each study team applied a number of different methods, including simple comparisons between treatment and control groups of changes in average consumption over time (“difference of differences”), ordinary least squares regression to estimate consumption changes associated with inclusion in the treatment group, and pooled time-series cross-sectional

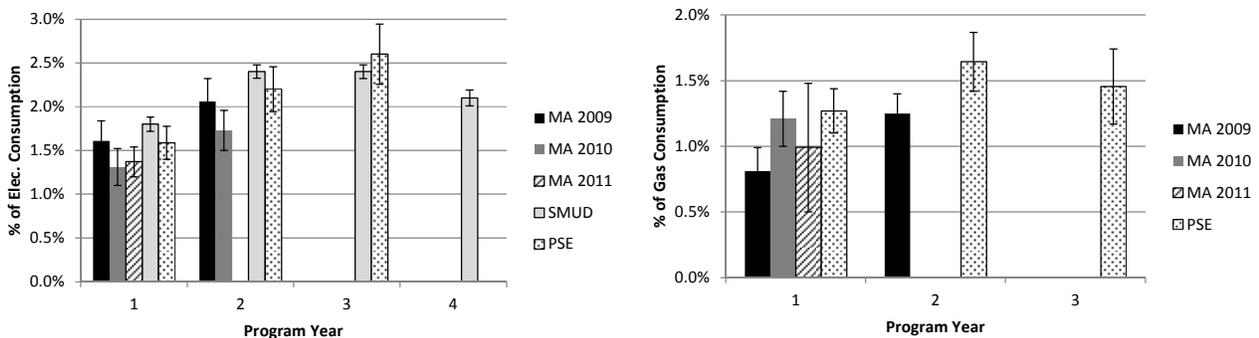
approach, which, at least theoretically, controls most effectively for the effect on the savings estimate of potential systematic differences between the treatment and control groups that may persist after random assignment. These pooled techniques are also referred to as panel or fixed effects regression and were used most often to represent savings. The methods used in these studies are generally consistent with those recommended by the State and Local Energy Efficiency Action Network, a project to support program administrators funded by the U. S. Department of Energy. (SEE Action, 2012)

- **Customer surveys.** All four studies included surveys of customers in the Treatment and Control groups, and focused on identifying the energy efficiency actions both groups took in the post-treatment period. Several also included on-site inspections to verify reported installations of energy efficiency measures.
- **Cross-participation analysis.** The PSE and MA studies included analyses of participation in other energy efficiency programs by customers in the Treatment and Control groups, using the cross-referencing of account numbers from the billing analysis to databases of participants in other programs.

Estimates of Annual Savings and Persistence of Savings

Figure 1 summarizes the key findings of the Massachusetts, SMUD, and PSE studies in regard to average annual savings associated with inclusion in the treatment group and their persistence over time. Average annual electric savings associated with first-time inclusion in the Treatment group ranged from 1.11 percent to 2.13 percent of pre-treatment use. The four programs with gas customers registered average first-year savings of 0.7 percent and 1.33 percent. The PG&E Gamma trial, which targeted a cross-section of customers achieved similar results. Given the large sample sizes, the confidence intervals around the average estimates were fairly narrow – ranging from 3 to 16 percent of the estimated electric savings. Gas confidence intervals, because of the greater seasonal variation, are generally wider ranging, from 12 to 49 percent.

Figure 1. Summary of Annual Electric and Gas Savings Estimates with Confidence Intervals



Sources: KEMA 2012, Opinion Dynamics Corp. and Navigant Consulting 2012, Integral Analytics 2012.

In energy terms, average first-year electric savings per participant ranged from 184 kWh to 241 kWh. To put those results in perspective, annual energy use reduction achieved through the installation of a CFL to replace an incandescent bulb is roughly 50 kWh per year. So, the annual electric savings associated with assignment to the Treatment group is equivalent to savings from replacing 3 – 4 incandescent bulbs with CFLs. Similarly, gas savings associated with assignment to the Treatment Group is roughly equivalent to the engineering-based estimate of savings achieved by

installing a faucet aerator to reduce hot water use (VEIC 2011).

Figure 1 also illustrates the change in energy savings over time for individual participation cohorts. In all cases for which at least two annual savings estimates are available, savings increase from the first to the second year. This result likely reflects the customers' learning curves as well as time required to implement efficiency measures. In two of the three series with three annual savings estimates, savings decline between Years 2 and 3, but remain higher than the first year savings. Average electric savings realized by the PSE treatment continued to grow, although the difference between Years 2 and 3 were relatively small. Average Year 4 electric savings among SMUD customers decreased from Year 3, but remained higher than the average savings in Year 1.

Customer Actions and Energy Savings

To assess the potential persistence of observed savings, we need to understand what measures customers installed and which behaviors they initiated as a result of exposure to the program. Moreover, to better assess the cost-effectiveness of comparative feedback programs, we need to consider whether all of the savings observed among the Treatment Group was achieved due to the influence of the feedback program, or whether some portion was due to participation in other incentive-based programs in the sponsors' portfolios.

Within the framework established by the experimental design, consumers in the Treatment group can generate savings through three mechanisms:

- ***Incremental participation in other efficiency programs.*** Customers in the Treatment group participate more frequently, earlier in the program cycle, and/or implement more measures through other efficiency programs offered by the feedback program sponsor than their counterparts in the Control group.
- ***Incremental installation of efficiency measures outside of other sponsor programs.*** Customers in the Treatment group install a greater number of energy efficiency measures on their own, without program assistance, than their counterparts in the Control group during the program period.
- ***Incremental adoption of efficiency and conservation behaviors.*** Customers in the Treatment group adopt efficiency and conservation behaviors (such as lowering thermostat settings or unplugging power supplies for small electronic appliances) to a greater extent than their counterparts in the Control group.

In this section, we review findings from the evaluations to assess the nature and magnitude of savings generated by these mechanisms.

Savings from incremental participation in other efficiency programs. All four studies reviewed here contained analyses of participation by members of the Control and Treatment groups in other energy efficiency programs offered by the feedback program sponsors. The evaluators used data merges to identify which Treatment and Control group members had participated in other "downstream" programs offered by the sponsors, both before and after the program period. The difference between the Treatment and Control groups in the pre/post program change in participation (Difference in Differences) was identified as the feedback program effect. This effect is sometimes referred to as "lift". The analysts for Massachusetts, PSE, and PG&E used information in the incentive program files to estimate annual savings from these measures and the date of installation. Savings were adjusted to reflect the portion of the evaluation period for which the measure was installed and, for measures that affected thermal end uses, the portion of annual heating and cooling loads included in the installation period. The SMUD program evaluators took a different approach to estimating savings associated with "lift". They included an interaction term in the fixed effects model for indicating feedback program participation, post-program period, and participation in another program *after* the feedback program started. They used the coefficient on that interaction variable to estimate incremental savings associated with participation in other programs.

Table 2 displays the methods and results of analysis of cross-program participation and savings from the different studies. Key observations from this table include the following.

- **Effect of treatment on participation rate.** Among the various sponsors, participation in other efficiency programs ranged from 2.5 – 4.0 percent among the treatment and control groups prior to feedback program launch. Incremental participation among the Treatment group ranged from 0.4 – 1.0 percent, or 10 – 25 percent of the baseline level. This is a significant boost.
- **Contribution of participation in other programs to total estimated savings.** All of the studies that used a measure-level engineering approach to estimate energy savings from incremental participation in other programs low savings, ranging from less than 1 percent of the observed change in consumption associated with assignment to the treatment group to around 10 percent. By contrast, the SMUD study, which used a modeling approach, found that nearly a third of total estimated savings were associated with participation in other programs. This finding may be an artifact of the modeling method, since voluntary program participation is associated with high levels of interest in energy efficiency, which in turn is likely associated with high levels of overall savings. Some portion of these savings should be subtracted from savings observed via the billing analysis to estimate the unique of the feedback program. Following regulatory advice, PSE allocates all joint savings to the incentive programs. The Massachusetts study deducts only a portion that reflects the incremental effect of assignment to the Treatment group on incentive program participation.

Table 2. Incremental Participation in Other Programs and Savings Associated with Assignment to the Treatment Group

Sponsor	Fuel	Cohort (Year)	Incremental Participation	Energy Savings from Other Programs: Calculation Method	Energy Saved/Yr	% of Total Est. Savings
National Grid MA	G	2009 (2)	0.37%	Measure-level engineering	--*	--*
National Grid MA	G	2011 (1)	0.68%	Measure-level engineering	--*	--*
National Grid MA	E	2009 (2)	0.25%	Measure-level engineering	--*	--*
National Grid MA	E	2010 (1)	0.85%	Measure-level engineering	7.3 Th	4.8%
National Grid MA	E	2010 (2)	0.26%	Measure-level engineering	12.0 kWh	6.1%
NSTAR MA	G	2011 (2)	0.97%	Measure-level engineering	0.4 kWh	2.3%
PSE	G	2011 (3)	0.04%	Measure-level engineering	1.3 Th	10%
PSE	E	2011 (3)	0.04%	Measure-level engineering	2.0 kWh	0.7%
PG&E	E	2011	0.70%	Measure-level engineering	0.7 kWh	0.7%
SMUD	E	Mixed	0.4%	Fixed-effect regression	79 kWh	33%

* - not significantly different from zero. For several MA cohorts not shown in the table, there was no statistically significant effect on participation.

Sources: KEMA 2012, Opinion Dynamics Corp. & Navigant Consulting 2012, Integral Analytics 2012.

Gauging the effect of feedback programs on CFL purchase and installation poses a further challenge in that most program sponsors use an upstream incentive approach whereby individual purchases of program-supported bulbs are not registered in a tracking database. In the case of PSE, evaluators estimated savings from purchase of CFLs separately as a program measure, using the results of the customer survey. Using a difference of differences approach, they estimated the effect of exposure to the feedback program on CFL installation at 0.15 bulbs per household. The PG&E study estimated incremental installations of 0.34 bulbs using data from an on-site socket inventory of sample and summary results of the most recent upstream lighting program evaluation.

Savings from incremental installation of measures without assistance from other

programs. Both the Massachusetts and PSE studies contained surveys of customers in the Treatment and Control groups designed to characterize the energy efficiency measures taken by those customers during the period of the experiment. Measures addressed by the survey included purchase of efficient appliances, physical improvements to the thermal shell, and changes in energy-related behaviors such as thermostat settings. The Massachusetts survey was fielded during the first year of the program. The PSE study first fielded a customer survey in the Program Year 3, and the sample included customers in an experimental group that had stopped receiving reports in the third year. This is referred to as the “Suspended” group. Table 3 displays data from the two studies on customer self-reports of energy-savings measures implemented during the study period by Treatment and Control groups, aggregated by major end-use technology group. The shaded cells indicate end-use technology groups in which the level of reported efficient technology adoption was significantly higher among the Treatment group.

Table 3. Percentage of Sample Control and Treatment Groups Reporting Efficient Measure Adoption in Program Period

Measure Category	National Grid (Electric)		National Grid (Gas)		National Grid (All Fuels)		Puget Sound Energy		
	Control	Treatment	Control	Treatment	Control	Treatment	Control	Treatment	Suspended
(n)	(251)	(250)	(250)	(251)	(501)	(501)	(502)	(452)	(498)
Heating/Cooling Equipmt.	8.6%	11.9%	8.1%	8.6%	8.4%	10.2%	2.7%	3.3%	2.1%
Appliances	22.8%	28.2%	16.8%	21.5%	19.8%	24.8%	10.0%	11.3%	12.6%
Consumer Electronics	14.0%	22.8%	13.2%	17.9%	13.6%	20.2%	31.5%	27.6%	38.8%
Water Heating	n/a	n/a	n/a	n/a	n/a	n/a	1.8%	3.5%	2.4%
Building Envelope	10.7%	18.0%	7.3%	13.9%	9.0%	16.0%	5.2%	5.4%	4.3%

Shaded cells represent statistically significant differences between control and treatment groups at $p > 90\%$
Sources: KEMA 2012, Opinion Dynamics Corporation and Navigant Consulting 2012.

The results summarized in Table 3 show that there is an inconsistent but, for some types of measures, significant association between assignment to the Treatment group and implementation of energy efficiency measures. None of the studies identified a significant difference between the Treatment and Control groups in the rate of implementation for efficient heating and cooling measures. Most of the larger differences observed involved relatively low-cost measures, such as purchase of efficient appliances and consumer electronics, which have modest unit savings levels. However, the Massachusetts surveys did detect a large difference in the fraction of households that implemented expensive building envelope measures, such as insulation and replacement windows: 16 percent of the Treatment Group versus 9 percent for the Control group.

Savings from incremental adoption of energy efficiency and conservation behaviors. The surveys conducted for all four studies reviewed asked respondents whether they had adopted or increased the frequency of a long list of energy efficiency and conservation practices in the prior year. Categories of practices questioned included thermostat settings for heating, cooling, and water heating equipment, HVAC and refrigerator maintenance, unplugging idle electronics, cold water washing, and so forth. In none of these measure categories did evaluators find any significant differences in the rate of adoption between Treatment and Control groups. In fact, in most cases, the adoption rates were nearly identical.

The lack of apparent differences between the Control and Treatment groups in practice adoption likely reflects the limitations of survey techniques as much as the underlying similarities between the groups. The billing analyses, supported by tens of thousands of observations were able to discern small differences in the groups on a single variable, namely annual consumption. In this case with 500 or fewer observations per group covering scores of variables, it is unsurprising that the

surveys were unable to identify significant differences in patterns of behavior.

The pattern of differences in monthly consumption between the Treatment and Control groups yields some insight into the behaviors that are likely to be driving savings. Analysis of monthly gas savings for PSE participants clearly showed that savings were much greater during the winter months, which suggests that most gas savings were being achieved through heating-related measures, most likely lowering of thermostat settings. By contrast, monthly savings of electricity were constant through the year, suggesting that most were generated through reductions in non-weather related end-uses such as lighting and plug loads.

Persistence of Observed Savings in the Treatment Groups after Program Suspension

Both PSE and SMUD have suspended the delivery of feedback reports to subsets of their Treatment groups to assess the effect of suspension on the persistence of savings. SMUD found that average monthly savings as a percentage of pre-program consumption decreased from 2.3 to 1.6 percent for a group of customers whose reports were suspended after Year 3 in the program. Savings decreased only from 2.3 to 2.1 percent for the group that continued to receive reports. PSE found that savings for the “suspended” group decreased from 2.3 to 1.8 percent from Year 2 to Year 3, while savings for the continuing group increased from 2.3 to 2.6 percent. Given that some of the net savings identified through the billing analysis is due to installation of long-lived measures, it is unlikely that savings will fall to zero in the medium term – say 4 – 7 years.

Effect of Differences in Report Frequency and Mode of Delivery

The random assignment capability inherent in the feedback report program model supports evaluation and comparison of the effectiveness of different implementation approaches as well as evaluation of overall program effects on consumption. The SMUD and PSE evaluations randomly assigned subsets of the Treatment group to receive reports quarterly versus monthly. Table 4 summarizes the comparison of savings between the groups that received monthly and quarterly reports during the first year of program operations. In all cases for which data are available, customers receiving monthly feedback reports achieved higher savings than those receiving quarterly reports. The difference for electric customers was relatively large, around 30 percent. This difference could affect cost-effectiveness of the larger program, depending on the split of fixed and variable program costs. The difference for gas customers was 17 percent. The PSE study also found that customers receiving quarterly reports did *not* achieve increased electric savings in successive analysis periods whereas those who received monthly savings did increase savings from one period to the next. This difference in savings patterns over time did not occur among gas customers. That is, customers who received quarterly reports, as well as those who received monthly reports, increased their gas savings over time.

Table 4. Savings by Treatment Groups with Quarterly v. Monthly Feedback

Sponsor/Fuel	Percent of Pre-Program Usage	
	Monthly	Quarterly
SMUD/Electric	2.3%	1.6%
PSE/Electric	1.9%	1.3%
PSE/Gas	1.2%	1.0%

SMUD experimented with a “seasonal burst” schedule of four monthly reports during the peak season of May through August. Customers in this subgroup reduced monthly energy use by 1.2

percent, versus 2.3 percent for the group as a whole.

SMUD also experimented with different report designs and modes of delivery. Researchers found that the graphic design of the paper report had no apparent effect on levels of savings achieved. In the most recent year of the program, SMUD delivered feedback reports via e-mail to a set of 5,930 customers who had enrolled in the company's on-line "Your Account" service for bill payment. The savings for these customers was considerably lower than recipients of paper reports. However, the evaluators noted that the control group chosen for comparison to the e-report recipients had much higher pre-program consumption than the treatment group, and that this mismatch may have compromised the results of the billing analysis.

Relationship of Pre-program Consumption Levels to Savings Achieved

All four of the studies assessed the relationship between customer attributes and levels of savings associated with assignment to the Treatment Group. The range of attributes analyzed included: the presence of a pool, spa, or electric heat; square footage of the home (available from assessor's records for SMUD); the age of the residence; assessed value of the home; and pre-treatment level of consumption relative to other customers in the treatment group. Of all of these attributes, only the relative pre-treatment consumption level was found to have a strong relationship to absolute kWh and percentage savings in all programs. As an example, the PSE Year 3 study (KEMA, 2012) addressed the issue of the relationship between pre-treatment consumption and savings levels by modeling savings for each pre-program year consumption quintile in relation only to the corresponding quintile in the Control group. Figure 3 shows savings as a percentage of pre-program weather normalized electric consumption by year and pre-program consumption quintile. In each year, the customers in the highest quintile (20 percent) ranked by pre-treatment weather normalized consumption accounted for roughly 50 percent of all program savings. The top two quintiles accounted for 75 percent of total program savings. These findings suggest that the cost effectiveness of the program could be increased by targeting customers with high levels of annual use relative to their peers. Except for the presence of spas, none of the other customer attributes examined showed statistically significant effects on energy savings associated with assignment to the Treatment group.

Figure 3. Percent Savings as a Percent Pre-Treatment Weather Normalized Consumption by Relative Pre-Treatment Consumption

Participant Characterizations of Response to Home Energy Reports

The Massachusetts and PSE studies incorporated surveys of customers in the Treatment groups that elicited information on their response to the home energy reports. The key findings from these surveys were as follows.

- **The majority of customers in the treatment group read and reviewed their home energy reports.** Ninety-two percent of sample Treatment group in the PSE territory recalled seeing the reports, as did 94 percent of the recipients in the Massachusetts sample. Among the Massachusetts customers, 59 percent reported that they read all of the reports they received; 29 percent reported reading most of them. In the PSE territory, 70 percent of respondents reported reading all of the reports; 14 percent reported read most of them.
- **Roughly one-half of home energy report recipients characterize them as useful.** The sample PSE customers were asked to rate how useful they found the home energy reports on a five point scale. Eighteen percent characterized them as “5 - Very useful”; 23 percent rated them as 4. The element of the report identified as most useful was the comparison of current monthly usage the respondent’s own usage in the same month of the prior year (48 percent rated “Very Useful”). Only 25 percent rated the comparison to neighbors as “Very Useful”. These findings might indicate a difference in the value customers accord to information that motivates action versus information that guides action. Among Massachusetts recipients, 41 percent found the reports useful for providing “new ways to save energy in [their] home[s]”.
- **Relatively few respondents identified a causal link between receipt of the reports and energy use reduction actions they had taken.** Among PSE report recipients, 37 percent claimed that the reports stimulated them to adopt new energy conservation habits; 29 percent claimed that the reports stimulated them to purchase energy efficient equipment. The Massachusetts study did not include questions on attribution in the telephone survey of report recipients. The study included set of in-home interviews with a small sample of customers (n=11), and the topic of causation was explored in those interviews. Only one of the respondents to the in-depth interviews drew a direct causal link between receipt of the reports and actions the household had taken to save energy.

Cost Effectiveness

One of the principal applications of evaluation is to guide efforts to reach savings goals and to make programs more cost-effective. To support this discussion, we have compiled in Table 5 planning estimates of program participation, energy savings, benefits, costs, and cost-effectiveness from the 2013 -2015 Energy Efficiency Plans submitted by the investor-owned energy efficiency Program Administrators (PAs) in Massachusetts. (Massachusetts EEAC) Of course, costs and savings will vary by jurisdiction, so the results shown in Table 5 cannot be regarded as universal. However, they have the benefit of being recent, publicly accessible, and based on three years of independently-evaluated program performance.

A number of key observations stand out on the Participation and Benefits side of Table 5. First, the 840,431 participants whom the PAs plan to enroll account for 36 percent of all their residential customers. Moreover, that figure represents a near tripling of 2012 enrollment. Clearly, the PAs are planning to scale up current feedback program operations significantly. Second, the PAs have made the conservative assumption that lifetime savings equal first-year annual savings. That is, they assume that savings would decrease to zero if the program were suspended. Evaluation research findings discussed above strongly suggest that some savings would persist for at least a year if not more. Finally, projected average savings per household are considerably below those realized by the

program in its first years. This may reflect an assumption that savings will decrease as the program expands and takes in a larger portion of customers with lower levels of consumption. Despite these conservative assumption, the modeled TRC Benefit-Cost ratio is fairly robust at 1.31, although considerably lower than the residential portfolio average.

Table 5. Selected Benefit and Cost Projections: Massachusetts Energy Efficiency Plan 2013 – 2015

Participation and Benefits		Costs and Cost Effectiveness	
Number of Participants	840,431	Costs – Total	\$9,050,839
Annual Energy Savings (MWh/Year)	99,733	Cost per Treatment Group Household	\$11.34
Lifetime Energy Savings (MWh)	99,733	Program Planning and Admin.	2%
Annual Savings/Participant (kWh/Year)	119	Marketing and Advertising	2%
Peak Demand Reduction (MW)	21	Participant Costs	73%
		Sales, Tech Assistance & Training	20%
Benefits – Total	\$9,534,393	Evaluation and Market Research	3%
Capacity	26%		
Energy	70%	TRC Benefit-Cost Ratio	1.31
Demand Red'n Induced Price Effects*	4%	TRC Benefit Cost Ratio All Residential	3.24

* Decreases in wholesale electric prices occasioned by demand reductions during high price periods.

On the cost side, it is interesting to note that participant costs – mainly the production and mailing of the feedback reports – account for 73 percent of the total program budget. Thus, strategies to limit those costs via changes in the frequency of mailings or the use of e-mail report delivery are worth at least worth trying. As SMUD’s experience suggests, it is important to think through evaluation issues in the construction of control groups as part of the implementation of any such operational experiments so that their effects can be accurately gauged.

Conclusions

As program administrators at scores of utilities across the U. S. and abroad prepare to launch feedback report programs, it is useful to consider the implications of the studies reviewed for program design and evaluation, and the strength of evidence behind those implications. First, the findings that can be stated with some certainty include the following:

- Customers assigned at random to receive feedback reports reduce their annual consumption by a measurable amount, usually in the range of 1 – 3 percent, depending on housing stock, climate conditions, and fuels.
- Savings persist and often grow through the second program year, and continue to be positive in the third year and fourth years.
- Comparative feedback programs are likely to be cost effective, even if conservative assumptions regarding the persistence of savings are applied.
- Only a small portion of the savings realized by the Treatment group appears to be related to incremental participation in other energy efficiency programs. Overlap of program effects needs to be accounted for in evaluation. Depending on assumptions used in the cost-effectiveness assessment, overlap with other programs may compromise the cost-effectiveness of the feedback report approach.
- Savings increase with frequency of reports.
- Savings persist if the frequency of reports is reduced or even if they are suspended entirely. However, the level of savings decreases under these circumstances.

Despite these encouraging early results, a number of uncertainties remain in regard to the value of feedback reports as their deployment expands. Future program design and evaluation work should address the following issues.

- **Identify savings mechanisms.** The mechanisms by which customers in the Treatment save energy remain unclear. Without a better understanding of customer actions in response to the feedback reports, it will be difficult to assess their persistence. To address this issue, evaluators and program administrators should consider fielding surveys of Treatment and Control Group members early and in each subsequent program year to track changes in behavior. These surveys should focus only on a few behaviors that early research suggests are key to realization of savings, such as installation of lighting and HVAC measures, control of lighting and plug loads, and thermostat control. Findings from evaluations of audit programs show that it customers often require a number of years to implement more costly measures. (Itron, 2008) Highly focused surveys will reduce respondent burden, improve response rates, and facilitate timely feedback into program design and management.
- **Identify mechanisms to increase average savings per customer.** This issue has been explored in the evaluations summarized here. Most recommendations from professional evaluators and customers themselves focus of providing more customer-specific information to report recipients to guide energy use reduction activities. Beyond changes in messaging and delivery mode, the major changes afoot in comparative feedback programs is their linkage with or integration into more intensive marketing and behavioral efforts. These include a variety of opt-in designs involving web portals, community-based outreach, use of premiums or rewards such as airline miles, and gamification, as well as more intentional targeting to narrow customer segments. (EnerNOC, 2013; Cisneros, 2012) Small scale implementations of these approaches suggest that they can produce savings per customer of 6 – 11 percent, in well in excess of the 1 – 3 percent observed for the reports with random assignment approach. (Ehrhardt-Martinez et al., 2010; Cole, 2012)
- **Apply the right evaluation tools for the job.** Given the expected low level of savings per customer from comparative feedback reports, random control trials (RCT) were a necessary first step to verify that the program approach does, in fact, produce net savings and that those savings do persist beyond the program year. However, RCT is not a panacea, even for the modest evaluation agenda initially established for these programs. One has only to read the fairly tortuous report sections on savings from participation in other programs to reach that conclusion. As comparative feedback programs evolve into more voluntary and targeted strategies, we will need to bring to bear on their evaluation the full repertoire of quasi-experimental methods such as discrete choice modeling and more straightforward analysis of billing and survey data. The availability of more frequent billing reads from AMI systems may facilitate such evaluations, as has been the case in small trials of home energy management systems (AECOM, 2009). Generally, however, designers and administrators of programs with voluntary elements will need to target savings higher than 1 – 3 percent if they expect evaluation of those savings to survive the familiar list of threats to validity.

References

AECOM. 2011. *Energy Demand Research Project: Final Analysis. Appendix D: SSE Community Trials*. London: Office of Gas and Electricity Markets (OFGEM).
<http://www.ofgem.gov.uk/sustainability/edrp/Documents1/EDRP%20Appendix%20D%20SSE%20community%20trials.pdf>

- Allcott, H. 2009. "Social Norms and Energy Conservation." *Journal of Public Economics*. 2011, doi:10.1016/j.jpubeco.2011.03.003
- Cisneros, Bruce. 2012. "New Insights for Home Energy Reports: Persistence, Targeting Effectiveness, and More" Presentation to the 2012 Behavior, Energy and Climate Conference. Sacramento, CA.
- Cole, Paul. 2012. "The Behavioral Model: A Case Study that Achieved 11 % Percent Persistent Savings in a Three Year Energy Monitoring Pilot", Presentation to the 2012 Behavior, Energy and Climate Conference. Sacramento, CA.
- Ehrhardt-Martinez , Karen, Kat A. Donnelly, and John A. "Skip" Laitner. 2010. *Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities*. Washington, DC: American Council for an Energy-Efficient Economy.
- EnerNOC Utility Solutions. 2013. *Paving the Way for a Richer Mix of Residential Behavior Programs*. The California Investor-Owned Utilities.
- Integral Analytics, Inc. 2012. *Impact & Persistence Evaluation Report: Sacramento Municipal Utility District Home Energy Report Program*. Sacramento, CA: Sacramento Municipal Utility District.
- KEMA, Inc. 2012. *Puget Sound Energy's Home Energy Reports Program: Impact and Process Evaluation*. Bellevue, WA: Puget Sound Energy.
- Massachusetts Energy Efficiency Advisory Council (EEAC), Three-Year Energy Efficiency Plan: Statewide Cost & Savings Tables. <http://www.ma-eeac.org/Three%20Year%20Plans.html>
- Itron, Inc. 2008. *Evaluation of the 2004-2005 Nonresidential Audit and PG&E Local Program*. San Francisco, CA: California Public Utilities Commission.
- Opinion Dynamics Corporation and Navigant Consulting, 2012. *Massachusetts Three-Year Cross-Cutting Behavioral Program Evaluation Integrated Report*. Waltham, MA: Massachusetts Energy Efficiency Advisory Council
- Sullivan, Michael, Brian Arthur Smith, Candice Churchwell. 2013. "Neighbor Comparisons Save Energy, but What Drives Savings?", *Proceedings of the 2013 International Energy Program Evaluation Conference*. Chicago, IL.
- State and Local Energy Efficiency Action Network. 2012. *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*. Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. <http://behavioranalytics.lbl.gov>.
- Vermont Energy Investment Corporation. 2011. *Mid-Atlantic Technical Reference Manual*. Lexington, MA: Northeast Energy Efficiency Partnerships, Inc.