Assessing Energy Savings Attributable to Home Energy Reports

Brian A. Smith, Pacific Gas and Electric Co. and Michael Sullivan, Freeman, Sullivan and Co.

ABSTRACT

The current energy efficiency ("EE") program cycle in California (2010-2012) is the first time that the California Public Utilities Commission (CPUC) is allowing California's investor-owned utilities (IOUs) to claim energy savings resulting from behavior-based efficiency programs. The decision authorizing utilities to claim savings from behavioral-based programs mandates that experimental design methodologies be used and explicitly points out that special care will be necessary "to ensure that savings credited to these programs do not represent double-counting" so the issue of attribution of savings is critical. Pacific Gas and Electric Company (PG&E) is the first California IOU seeking to claim savings from this type of program. This paper describes the requirements for demonstrating energy savings from behavioral programs, outlines the research design that will be used to determine the magnitude of savings resulting from a specific behavioral program based on home energy reports ("reports") and discusses how the important issues related to attribution of savings between multiple, competing programs can be resolved. We show that, as long as the evaluation is designed to compare two populations which in statistical terms are in no way different except for the treatment of the program, the measured savings should be those attributable to the program, provided the experiment is properly designed. A key challenge in the design of this behavior-based experiment is to demonstrate that this behavior-based program provides added value to the many efficiency programs already underway, and that the electric and gas savings claimed will stand up to heavy scrutiny by the impact evaluation to be managed by the CPUC.

Introduction

This paper describes the experimental design for PG&E's test of home energy reports ("reports"). Treatments, assignment strategy, stratification, sample design and analysis framework (i.e., panel regression) are presented. In addition, this paper attempts to advance the use of measurement procedures to isolate impacts of the program from those of other EE programs (e.g., lighting, other rebate programs) so that double-counting of savings is avoided.

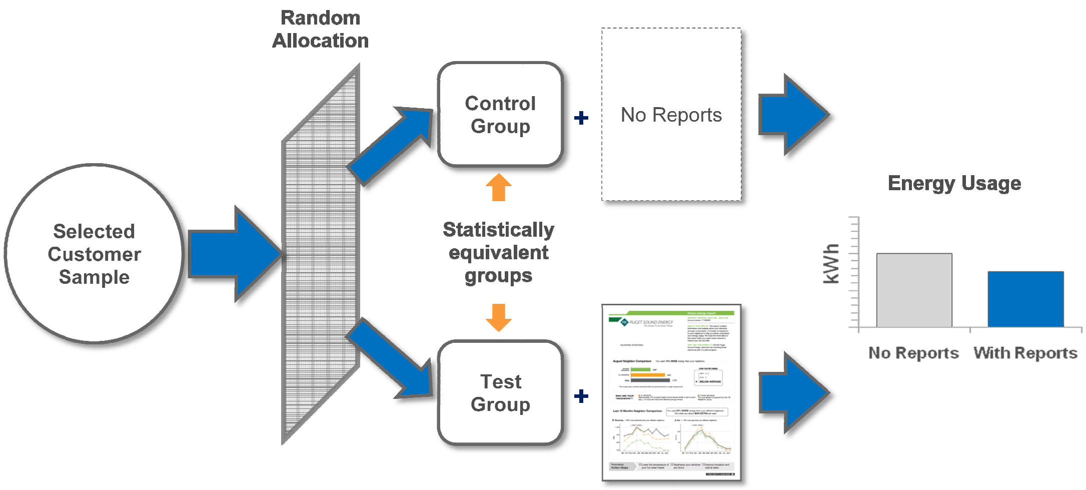
The general idea of reports[1] is that, by presenting comparative energy use feedback along with highly targeted, specific tips, consumers will make simple changes to their behavior (e.g., shutting off lights when they leave a room, adjusting thermostats, closing blinds, installing CFLs, buying more efficient appliances) that result, in aggregate, in large energy savings. The reports contain:

- A graphical representation of historical energy usage and variation (throughout the year, month, week, and day)
- Comparisons of energy use to a peer group (neighbors matched by residence size, building type, and proximity)
- General and personalized tips to reduce overall energy usage and/or peak usage, and
- The estimated savings benefit in dollars tied to specific energy savings behaviors.

---

[1] There are many ways that consumers can receive energy use information (e.g., similar feedback via email or web-based portals, direct feedback from monitors or sensors) and savings tips (e.g., general marketing campaigns, utility mailings and websites). This paper focuses on a test using an experimental design of mail-based reports produced by a private firm called OPOWER containing neighbor energy use feedback.

The cornerstone of experimental design is the use of random assignment of sample frames to treatment and control conditions of sufficient size so that differences between them can be measured. Experimental design is typical in much of medical and social science research (for example, in testing the effects of pharmaceutical drugs or the effects of a public health campaign) but is relatively new in the EE field. The beauty of experimental design is its ability to isolate the impacts of the treatment from all other impacts (e.g., exposure to general information campaigns or peripheral information, propensity to save energy, etc.). In this application of experimental design, households in the treatment groups will receive reports by mail while households in the control groups will not. Aside from getting the reports, households in treatment conditions will not be treated any differently than those in the control conditions. That is, there will be no difference between customers in treatment, control, and non-participating customers with respect to customer communications, exposure to ad campaigns, use of energy use feedback systems, or any aspect of utility service delivery. Since the treatment and control groups are essentially equal, there is no need to separate the effects of other measures or behavioral influencers which may cause changes in behavior. Any differences observed in energy use between the control and the test groups are attributable to the impact of the treatment. The following illustration shows how the experiment will be conducted.



The current EE program cycle in California (2010-2012, likely to be extended to a four-year cycle, 2010-2013) is the first that the California Public Utilities Commission (CPUC) is allowing California's investor-owned utilities (IOUs) to claim energy savings driven by behavior-based efficiency programs.[2] The CPUC's definition of "behavior-based efficiency programs" is fairly narrow, being limited to those that provide "information to residential subscribers relative to the amount of energy used by the metered residence compared to similar residences in the subscriber's geographical area." The decision allowing savings claims from behavior-based programs mandates that experimental design methodologies, as contained within the California Evaluation Protocols, be used. The decision explicitly states that special care will be necessary "to ensure that savings credited to these programs do not represent double-counting." Pacific Gas and Electric Company (PG&E) is the first California IOU seeking to claim savings from this type of program.

[2] Decision (D.) 10-04-029 provides for savings claims for behavior-based programs based on experimental design (pp. 40-42, see Conclusion of Law 18; OP 13). The policy determination in D.09-09-047 commits to crediting ex post savings for behavior programs in the 2010-2012 program cycle (D.10-04-029 OP 14).

Energy Savings Potential: Conventional vs. Behavioral Programs

As is typical throughout the industry, residential EE programs at PG&E tend to be "measure-based" whereby households receive efficient products discounted by retailers through "utility buy-downs (e.g., compact fluorescent lamps (CFLs)) or are encouraged to buy more efficient models through direct-to-consumer rebates (e.g., appliances). Rebates on CFLs are responsible for the largest share of energy savings of any residential efficiency measure administered by California IOUs. Rebated CFLs accounted for over half (56%) of the net kWh savings in the statewide portfolio for the 2006-2008 program cycle representing net annual energy impacts of 991,965,497 kWh from rebated CFLs used in residential settings in California.[3] There is disagreement whether residential lighting has been transformed[4] but as of 2010 only one in five standard light sockets are filled with CFLs in California residences. Still, many parties are pressing IOUs to identify new sources for residential energy savings.

California IOUs are under significant pressure to diversify their residential EE portfolios beyond the current measure-based approach for other reasons as well. California has set ambitious goals of reaching all 13 million existing homes with comprehensive EE improvements by 2020.[5] To achieve significant progress toward this goal, program efforts must be scaled significantly over the next nine years. Measures such as appliance rebates and appliance recycling are fraught with limitations including low penetration, high levels of free ridership and spillover, and rebate redemption breakage. It is time for a change of direction.

Behavior-based programs such as home energy reports may represent a significant opportunity to generate energy savings—and to fully exploit the energy savings made possible by the purchase of rebated widgets such as CFLs by using them effectively. Deployments of comparative energy use programs by a variety of utilities throughout the U.S. have yielded savings on the order of 1.5 to 3.5 percent. Experiments using Randomized Controlled Trials (RCT), with relatively large sample sizes have been used to detect changes on the order of 1.1% - 2.8 %.[6]

PG&E has engaged OPOWER, a privately-held company based in Arlington, VA, to produce customized reports. Competitors to OPOWER include Efficiency 2.0, an LLC based in New York City. The reports being produced for PG&E are similar to those that OPOWER produces for other utilities, but some content modules are specific to PG&E. OPOWER's report modules are tested over its many customers so that only the most effective ones are retained and further refined. From a program cost perspective, initial development costs are substantial, but per-respondent costs decline as the number of households included in the program increases. The bottom line is that the program size can scale well if energy savings are delivered as forecast. A sample report created by OPOWER for PG&E is shown in the illustration below.

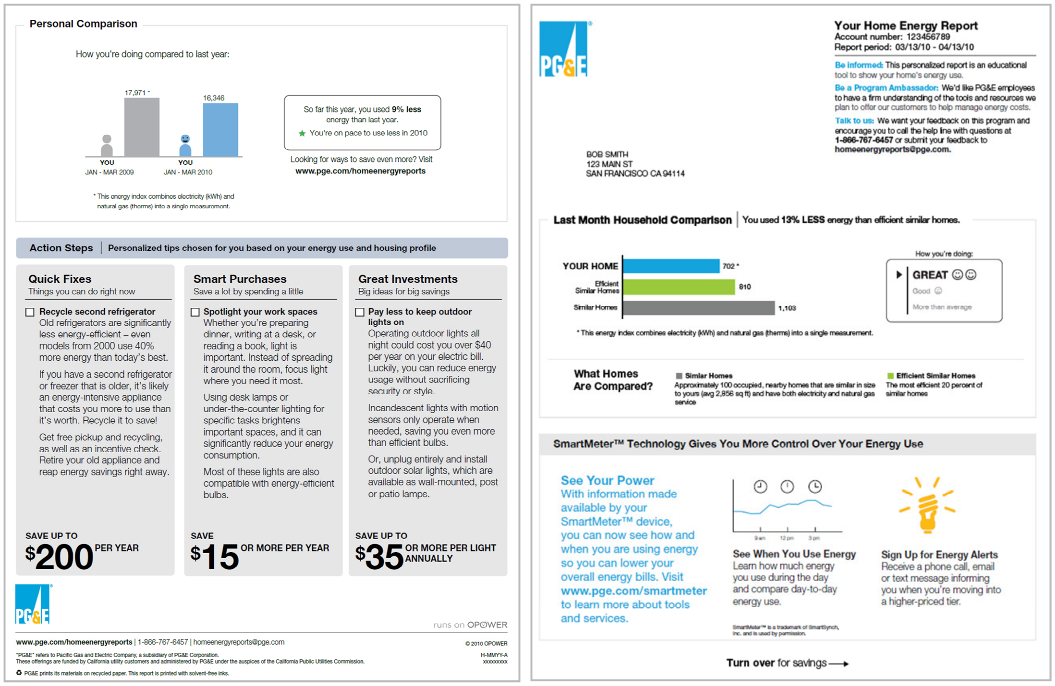
---

[3] See Final Evaluation Report: Upstream Lighting Program, February 8, 2010, Volume 1, pg i, Prepared by KEMA for the CPUC, available at http://calmac.org/, CALMAC Study ID: CPU0015.02.

[4] For a review see Market Transformation and Resource Acquisition: Challenges and Opportunities in California's Residential Efficiency Lighting Programs by Lara Ettenson and Noah Long, Natural Resources Defense Council, available at http://eec.ucdavis.edu/ACEEE/2010/data/papers/2106.pdf.

[5] See the CPUC's California Energy Efficiency Strategic Plan available at http://www.cpuc.ca.gov/NR/rdonlyres/A54B59C2-D571-440D-9477-3363726F573A/0/CAEnergyEfficiencyStrategicPlan_Jan2011.pdf.

[6] Some recent independent verifications of the effectiveness of home energy reports as implemented by OPOWER can be found here: http://www.opower.com/Results/IndependentVerification.aspx. The Consortium for Energy Efficiency (CEE) tracks utilities with similar programs for CEE members. See http://www.cee1.org/ for more information.



The test of home energy reports at PG&E will be conducted over several waves:

- An initial pre-test to approximately 10,000 households occupied by PG&E employees and retirees was launched in spring of 2011. The purpose of this test is to familiarize employees with report content and assess operational readiness to field information requests from customers receiving the reports in subsequent waves.
- The first test to approximately 60,000 households in the highest quartile of energy use will be launched in summer of 2011. The key purpose of this test is to estimate the energy savings potential for high energy usage customers.
- A second test to approximately 180,000 households representative of most PG&E residential customers (in terms of climate zone and energy use) will begin in autumn of 2011. The key purpose of this test is to estimate change in energy consumption resulting from exposure to the reports for a wide range of PG&E residential customers with different characteristics.
- A larger roll-out to as many as one million households will begin in 2012 once the preliminary results of the first two tests are available to inform targeting.

Given the scheduled roll-out of this program from 2011 to 2012, over one million PG&E households may be receiving the reports in 2013—providing that the tests are as successful at affecting energy savings as they have been in other jurisdictions. The following table displays estimated electric savings for the reports for 2013 and other residential electric efficiency measures already underway. Note that the estimates are independent: the estimated electric savings for the reports are in addition to the other efficiency measures shown as double-counting will be avoided. If report content persuades household recipients (treatment group) to take part in other utility-sponsored efficiency programs (such as appliance rebates as determined by a study of utility records or purchase of discounted CFLs as determined by post hoc surveys) at a rate that is greater than households in the control group, the resulting energy savings will be credited to the other utility programs and not to the home energy reports so that double-counting of energy savings is avoided.

ESTIMATED ELECTRIC SAVINGS FOR 2013	EE Program Gross kWh Savings	% of Portfolio Savings Goal
HOME ENERGY REPORTS	231,511,000	39%
BASIC LIGHTING	116,862,651	20%
ADVANCED CONSUMER LIGHTING	72,605,477	12%
BCE CONSUMER ELECTRONICS	62,948,520	11%
HOME EE REBATES	52,732,185	9%
APPLIANCE RECYCLING	29,336,475	5%
MULTI-FAMILY REBATES	21,115,719	4%
IN-HOME ENERGY AUDITS	8,458,590	1%
TOTAL RESIDENTIAL PORTFOLIO	595,570,616	100%

Since PG&E provides both electricity and natural gas to a majority of its residential customers, the potential of the reports to encourage gas savings is of equal importance. Moreover, as a result of measure interaction[7] as calculated by the CPUC, PG&E's attainment of its electrical goals is made more difficult due to offsets made for interactive effects. The following table displays estimated natural gas savings for the reports for 2013 and other residential gas efficiency measures already underway. Note that the estimates are independent: the estimated natural gas savings for the reports are in addition to the other efficiency measures shown. As the table makes apparent, the reports have the potential of being a large source of natural gas savings.

ESTIMATED NATURAL GAS SAVINGS FOR 2013	EE Program Gross Therms Savings
HOME ENERGY REPORTS	6,427,762
HOME EE REBATES	3,291,104
IN-HOME ENERGY AUDITS	-
MULTI-FAMILY REBATES	-
APPLIANCE RECYCLING	(735,001)
BCE CONSUMER ELECTRONICS	(1,023,709)
ADVANCED CONSUMER LIGHTING	(1,286,870)
BASIC LIGHTING	(1,328,195)
TOTAL RES PORTFOLIO	5,345,092

Critical Measurement and Evaluation Requirements for Behavioral Programs

PG&E will be seeking to recover its program costs as well as earn substantial energy-saving shareholder incentives based on the response to customers to the reports. To do so, PG&E must conclusively demonstrate that:

[7] Measure interaction, also known as interactive effects, are estimated by the CPUC to account for the possibility that the installation of an energy efficiency measure that has a primary effect on one end use (such as CFLs being substituted for incandescent lamps for lighting) produces a corresponding secondary effect on another end use (such as increased use of natural gas for heating to compensate for the absence of the heat thrown off by incandescent lamps). See Brunner, Eric, Ford, Peter, McNulty, Mark, and Thayer, Mark: Compact fluorescent lighting and residential natural gas consumption: Testing for interactive effects in Energy Policy, Volume 38, Issue 3, March 2010 (pp. 1288-1296) for a recent study that questions their existence.

- Customers exposed to the reports reduced their energy use solely as a result of the reports and within very tight measurement tolerances. That is, they must demonstrate a reduction in energy consumption for households exposed to the reports within a statistically tight error band.
- The savings reported for the reports and the other energy-saving programs can be isolated from one another in such a way as to remove the possibility of over-counting for savings when cost recovery and savings incentives are calculated. In other words, PG&E must demonstrate that it is not claiming savings for the reports that it has claimed for other efficiency measures.

Here comes the issue of double-counting: some of the savings found in the households receiving the reports may be the result of energy-saving purchases that were already incented by other PG&E programs. For example, households receiving the reports may be more likely to have purchased an energy-efficient appliance for which PG&E pays rebates directly to residential customers. Or households receiving the reports may be more likely to have installed CFLs for which PG&E provided retailers a subsidy to lower the retail price. Utility records (for rebates paid "downstream" directly to residential customers) and post hoc surveys (for rebates paid "upstream" to retailers) will determine the difference between treatment and control groups in the uptake of incented measures. Comparison between the treatment and control groups will determine whether some of the energy savings found in the treatment group and not in the control group is a result of other utility EE programs. Energy savings due to increased EE program participation in the treatment groups will be attributed to the other programs and subtracted from the calculation of energy savings due to the reports, thereby avoiding double-counting. This approach has been used at other utilities, and the unique impact of the reports in several jurisdictions in affecting energy savings has been found to vary from 1.2% to 2.8% for electricity and from 1% to 1.5% for natural gas.

Overview of Possible Impacts of Behavioral Programs

In thinking about the impacts of behavior on household energy use there are two important classes of behavior to consider:

- Usage behaviors – Behaviors that affect the day to day use of appliances (e.g., thermostat settings, habits involving ventilation, lighting use, habits related to the use of unattended appliances like household computers and entertainment systems, use of power strips and timers to control "vampire" plug load etc.), and
- Appliance acquisition and shell improvement related behaviors – Behaviors that result in the replacement or acquisition of energy-using appliances or building shell improvements.

Most of the funding and effort expended to influence household energy use in the past three decades has been channeled through programs designed to influence the replacement of existing appliance, lighting and building shell designs with more efficient ones; or to cause the choice of more energy efficient alternatives when new appliances and building shells are acquired. In many states, utilities and third-party EE programs have spent literally billions of dollars on programs designed to achieve such energy efficiency gains.

In the beginning, the gains achieved by conventional EE programs were relatively easy to demonstrate because improvements in energy efficiency were calculated using fairly straightforward engineering calculations applied to appliance stock. However, with the passage of time these calculations have become increasingly controversial as evaluators and policy makers have realized that

that efficiency gains from EE programs targeted at technology improvements are occurring naturally in the market as a result of technical innovation and increasingly strict codes and standards for equipment and building shell performance. The result are net-to-gross (NTG) calculations that adjust savings claims to account for those who would have adopted the efficiency measures in the absence of utility rebates (free-ridership). A further and countervailing problem is this: EE programs can cause changes in energy efficiency that are not directly attributable to a particular upgrade incented by an EE program. For example, a household may purchase an energy-efficient appliance and receive a utility rebate, and a neighbor may decide to purchase a similar appliance but not receive a utility rebate. So, the measurement of the impacts of programs on acquisition behavior has become an increasingly difficult and controversial aspect of energy efficiency measurement and evaluation. In essence, consumer behavior has become a progressively more important issue in understanding the impacts of conventional EE programs. The big arguments about the efficacy of EE programs are not about the performance of technology. They are now about acquisition behavior. One often-heard debate is this: given the "greening" of energy consumption consciousness, would consumers have purchased the more efficient technology in the absence of the utility-sponsored program?

In contrast, until relatively recently, efforts to alter usage-related behaviors *per se* have received only meager attention. The historical lack of interest in altering usage-related behavior stemmed from several considerations surrounding the development of early EE initiatives during the late 1970s:

- Unlike the engineering and economic disciplines that were able to provide ample theory and empirical evidence that changing the stock of appliances and buildings would significantly alter energy use, the theories in the social sciences that could be used to develop reasonable approaches for causing changes in energy consumption were relatively primitive and, more importantly, empirically untested.
- Unlike the efficiency gains that arise from replacing appliances and improving building shell performance, it is relatively difficult to conclusively demonstrate that a given EE program intervention has caused a change in usage behavior – and subsequent change in energy consumption. Behavior changes result in subtle differences spread out over potentially large populations. The scale of study required to demonstrate changes in energy consumption arising from social interventions is more challenging (usually orders of magnitude so) to carry out than the engineering studies required to demonstrate energy efficiency improvements.
- Because human behavior is inherently changeable (in sometimes unpredictable ways), it was believed that changes in usage resulting from behavior change were likely to be temporary and easily reversible (particularly when compared to changes in the energy using infrastructure).
- Changes in usage related behavior were equated with reductions in the standard of living for households (e.g., increased discomfort from setting thermostats at uncomfortable levels); and energy policy programs that could be characterized in this manner were thought to be inherently unpopular with the public and correspondingly were unpopular with energy policy makers.

There have been several important changes in recent years that have resulted in an increased interest in behavior change on the part of policy makers:

- Improving the efficiency of appliances and building stock has proved to be more difficult to accomplish using a wide variety of seemingly reasonable interventions based on engineering and economic approaches (i.e., providing consumer education and then, when that was ineffective, economic incentives).

- At the same time, the need to reduce household energy consumption has become more pressing as the external (societal) costs of using fossil fuels (i.e., climate change, national security issues) have become more evident. In recent years, the technology for measuring and controlling household energy consumption has begun to evolve very quickly with the increasing market penetration of internet-based measurement and appliance control devices and advanced metering systems.
- Finally, in the past two years, several relatively robust experiments[8] designed to detect the impacts of feedback on energy consumption have demonstrated that household energy consumption can be reduced significantly using relatively simple and inexpensive procedures for providing feedback to households concerning their energy consumption in comparison with other "similar" households such as the reports being discussed here.

Challenges in Measuring the Impact of Home Energy Reports

Home energy reports provide households with a comparison of their energy use with that of a composite of "other" similar households on a monthly or quarterly basis in a direct mail communication mailed separately from the customers' utility bill. The reports also include tips for lowering energy use. We note that there are other ways for customers to receive energy use feedback (such as near real-time access to usage information using home displays made possible through automated metering infrastructure (AMI)). Distinguishing features of the reports being deployed in this experiment include the use of neighbor usage feedback, the distinct design of the reports and the mail-based delivery method.

Experiments using Randomized Controlled Trials (RCT) with the reports provided by OPOWER, using relatively large sample sizes, have detected changes on the order of 1.1% - 2.8 %. Near the upper end of this range, such programs are massively cost-effective. Near the lower end, they are marginal. Despite the progress that has been made in developing this apparently effective approach to changing household energy use behavior, there remains considerable uncertainty about several critical aspects of these programs. The test of the impacts of the reports must take account of these sources of complexity to provide useful information that can guide the development of a future cost effective program. The important issues are:

1. **Experimental procedures.** Minor differences in experimental procedures used in the publically-reported studies create uncertainty about the likely magnitude of the impacts that can be achieved for particular utilities. For example, some of the studies randomly selected neighborhoods for study (i.e., instead of individual households), while others excluded subjects from study groups when less than a certain number of comparable households were available within a two-mile radius to identify "composite comparison households" to be presented to treatment groups. Experimental differences in the ways in which the households were targeted, how neighbor energy comparisons were calculated, and the way the information was presented undoubtedly account for the wide range of impacts that have been observed to date. Understanding the drivers of savings magnitude will be critical for informing policy for compensation and incentive payments for individual utilities. Until the drivers are better understood it will be necessary for individual utilities to demonstrate the impacts of their specific programs.

[8] For a review of behavioral studies, see Freidrich, Katherine, Amman, Jennifer, Vaidyanathan, Shruti, and R. Neal Elliott. 2010. "Visible and Concrete Savings: Case Studies of Effective Behavioral Approaches to Improving Customer Energy Efficiency." Washington, D.C. Report E108, American Council for an Energy Efficient Economy. For a discussion of early experiments using OPOWER stimuli, see Allcott, Hunt. 2010. "Social Norms and Energy Conservation." Available from www.opower.com.

2. **Frequency of delivery.** The studies that have compared the impacts of different delivery frequencies have shown mixed results. While monthly delivery tends to generate more savings than lesser frequencies, monthly delivery also results in higher costs and recipient opt-out rates. We note that there are other ways to deliver similar information (e.g., email, web presentment), but these alternate forms of delivery are not the subject of this experiment.

3. **Persistence of effects.** A key unanswered question is whether ongoing delivery of the reports may eventually dull their effects. More tests are necessary to identify an optimal delivery strategy that maximizes customer engagement at minimum cost. The earliest OPOWER pilot programs were started in 2008 and some were conducted for only one year. Even the longest studies have only been in existence for two years and while the initial findings from these studies suggest that the impacts persist (and perhaps grow larger) over time, it is simply too soon to reach a conclusion on this issue.

4. **Correlation of energy usage with energy savings.** Not surprisingly, analysis of the results from prior studies indicates that the absolute magnitude of savings associated with the reports is correlated with the magnitude of household energy use: the more energy a household uses, the more savings they achieve as a result of exposure to the reports. In effect, people who use more have greater potential for savings, both in terms of the percentage of their energy use and in terms of the energy saved. Each report delivery comes at a cost and ensuring that the reports are maximally cost effective requires careful study of the relationship between energy usage and customer characteristics (e.g., climate zone, building type). It doesn't make economic sense to provide the reports to customers who cannot provide significant savings: likely those who already use less energy and those who are efficient but simply consume more due to lifestyle preferences such as hobbies and medical need that require more things in the home that require power (e.g., pool and aquarium pumps and heaters, specialty lights, large plasma TVs, computer and medical equipment, etc.) that households are unwilling and/or unable to do without. Cost aside, there are equity issues to consider: what if the reports produce less energy savings overall for low energy users (that is, less kWh and Therms consumed) but at a greater percentage reduction than for high energy users? Research will inform policymaking as these types of programs evolve in the near future (e.g., content refinement, delivery platform, frequency of report delivery).

5. **Interactive effects with other energy efficiency programs.** Efforts to isolate the effects of the reports from those of other EE programs in the same market have generally shown that these reports have modest influence on the uptake of EE program measures that are tracked by utility measure tracking systems. However, the impact of these programs on measures that are not directly tracked in IOU measure tracking systems (e.g., CFLs, timers, weather stripping, etc.) has been observed by surveying treatment and control group customers after the fact of exposure to the program. The data collection methods used in such studies (telephone interviewing) can be unreliable and therefore there is considerable uncertainty about the impact of the reports on these measures. Interactive effects with general efficiency marketing campaigns have not been assessed to our knowledge; very large sample sizes would be required to isolate such effects given their likely modest size.

PG&E's Measurement and Evaluation Design for Home Energy Reports

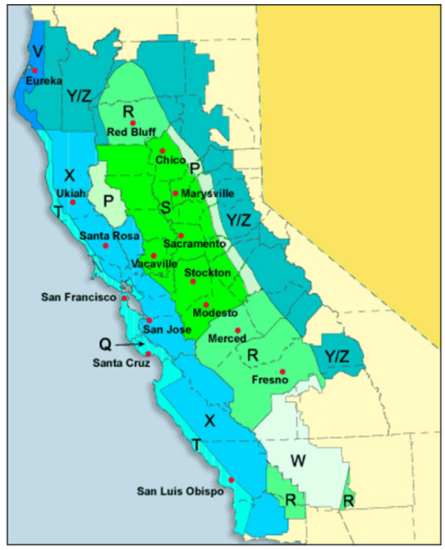
PG&E is seeking to develop the most cost-effective implementation of home energy reports that is possible given the characteristics of its diverse residential customer population. The cost effectiveness of the implementation will be determined by the amount of electricity and natural gas saved by the treatment groups as compared to the control groups. We expect that household characteristics, historical energy usage, and the frequency with which the report stimulus is transmitted to customers will drive the magnitude of the energy savings achieved. Finding the ideal combination of responsive customer segments and delivery frequency is the central challenge in designing this program. The expected key drivers of impact are discussed below:

1. **Target Customer Impact.** PG&E serves over 5 million households in Northern California, a region with extreme differences in climate. On most summer days the temperatures in the service territory vary from about 65 degrees F at the coast (where a very significant fraction of the population lives) to over 100 degrees F in the inland valleys. So the average impact of the reports may vary dramatically across the service territory, providing substantial savings in some high energy use locations (such as the Central Valley) and little or nothing in others (such as temperate coastal areas with low heating and cooling loads).
2. **Report Frequency.** Based on its experience, its home energy report provider (OPOWER) has recommended that PG&E adopt a customer contact protocol in which communication with the customer commences with three consecutive monthly reports followed by a new message every other month thereafter. This results in 6 total messages per customer per year (after the first year) with a set variable cost per customer per year. It may be possible to lower the frequency of contact with customers to quarterly after the first year and obtain similar results. Prior tests concerning the impacts of quarterly frequency are inconclusive owing to methodological problems with earlier tests. So PG&E will test the impact of quarterly frequency on energy savings in its evaluation.
3. **Commodity type.** PG&E provides different combinations of commodities across its service territory. In some areas it provides gas and electricity service. In other areas (i.e., those served by municipal electric utilities and coops) it provides only gas service. In other (sparsely populated) areas it supplies only electricity. This variation in the combinations of commodities that are provided creates complexity in program administration because different stimulus packages (reports) and delivery schedules must be created. More importantly, it produces a situation in which the report content and delivery schedules are not common to all customers within otherwise comparable treatment conditions. The necessary variations in the messages presented in the reports (e.g., tips and energy savings messages) to reflect customer fuel type differences in the reports add to the analytical complexity of the study.

Measuring the Energy Savings from Home Energy Reports

To measure the energy savings obtained from the reports and separate the impacts arising from changes in usage (behavioral) from those arising from appliance acquisition or building shell changes, PG&E will employ an randomized controlled trial (RCT) using the experimental design as described below. In the experiment, customers will be randomly selected from 8 different market segments: combinations of commodity type and baseline territory (these territories are similar to weather zone and are identified by single letters as shown in the map below). Within each group, households will be randomly assigned to one of three experimental conditions (standard frequency, reduced frequency and control groups).

About 15,000 customers will be assigned to each experimental condition – resulting in 180,000 treatment customers and 120,000 controls. Adjusting for estimated attrition of 50% over the course of the experiment that is planned to last two years, the resulting sample sizes per treatment and control condition are 10,000. The sample sizes for the treatment cells were identified using bootstrap sample design techniques to identify the minimum number of observations required to estimate parameters in a fixed effects panel regression model such that the parameters in the model will be within +/- 0.5% of the true population parameters 90% of the time – assuming the true population parameter is 1.0% and the resulting sample sizes will be 10,000 per cell.



The design will allow detection of differences in energy savings across geography and for different commodities as well as allow detection of minor differences in energy savings arising from the different contact frequency conditions. Importantly, the sample sizes are also large enough across the study to allow for meaningful analysis of the impacts of electricity consumption, household income, low-income program participation, and other potential stratification variables on energy savings.

The number of observations drawn from each kWh stratum will be equal, but it will not be proportional to the distribution of customers in the population. That is, the fourth quartile will over-represent high usage customers relative to what would occur in a simple random sample, and the first quartile will under-represent the distribution of low usage customers relative to what would occur in a simple random sample. This technique will improve the statistical reliability of estimates dramatically by increasing the number of observations that are taken in the highly variable upper strata. It will make the estimates less susceptible to "leverage" that very large customers can have on regression estimates when they occur in relatively small numbers. It will also tend to bias the sample in the direction of high users where savings are likely to occur—without underestimating the effects that low usage customers may be able to provide in the future. An illustration of the experimental design is presented below:

Service Type	Dual Fuel Customers																Electric Only Customers												Gas Only Customers			
Climate Band	R				S				X				W				R				S + P				X				Non-SMUD Gas-Only (Multiple Climate Zones)			
kWh Quartile	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Sample Sizes (000)	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75	3.75
Standard Frequency (000)	Treatment 1 (n=15)				Treatment 3 (n=15)				Treatment 5 (n=15)				Treatment 7 (n=15)				Treatment 9 (n=15)				Treatment 10 (n=15)				Treatment 11 (n=15)				Treatment 12 (n=15)			
Reduced Frequency (000)	Treatment 2 (n=15)				Treatment 4 (n=15)				Treatment 6 (n=15)				Treatment 8 (n=15)				-				-				-				-			
Control (000)	Control 1 (n=15)				Control 2 (n=15)				Control 3 (n=15)				Control 4 (n=15)				Control 5 (n=15)				Control 6 (n=15)				Control 7 (n=15)				Control 8 (n=15)			

Partitioning Direct and Indirect Effects

The use of the randomized controlled trial (RCT) design provides a relatively simple procedure for partitioning the effects of the reports from those of other utility EE programs. Because customers will be randomly assigned to treatment and control conditions within market segments, any difference between treatment and control groups in the participation rate for other PG&E EE measures will be a direct result of the treatment and will be measured using utility records and through on-site surveys. The calculation for partitioning report effects on the uptake of other utility rebates is straightforward for EE programs whose participants are directly tracked in utility databases: simply observe any differences in the rate of participation between treatment and control groups. If the rate of EE measure adoption in the treatment groups exceeds the rate in the control group, then this additional uptake will be subtracted from the total energy savings for the reports so that these savings are not counted twice. For example, if 5% of control group households, and 10% of the treatment group households, receive an appliance rebate, then the energy savings claimed for the reports will be reduced so that the additional energy savings found in the treatment group is attributed to the appliance rebate measure and not to the reports.

Isolation of the unique effects of measure-based utility EE programs which do not track end-user participants is more difficult, and potentially subject to greater measurement error. There are a number of so-called "upstream" programs in which utilities buy down the price of energy-efficient products to manufacturers or retailers to encourage changes in manufacturing, retailer stocking, and consumer purchase behaviors. CFLs, consumer electronics and home weatherization programs are examples of such measures at PG&E. Consumer rebates are not generally part of these programs, so sales of these measures are not tracked for specific end-users. Consequently there is no way to consult utility databases to observe the differences in the uptake of these products.

For such programs, it is necessary to directly study the differences in the uptake of relevant measures using statistical surveys of treatment and control group customers. It is obvious that survey errors (arising from coverage bias and survey non-response bias) are potentially serious sources of measurement error using this approach. Commonly-used survey measurement techniques (such as telephone interviewing) can be expected to produce response rates in the range of 25-35% using the most aggressive measurement protocols available to the industry. Even the best mail–based survey measurement techniques are unlikely to produce response rates in excess of 70% using the best available survey measurement protocols. Equally as serious will be challenges with sample size: large samples from both treatment and control groups will be required to have sufficient statistical power to detect significant differences in uptake of measures. Publicly-available evaluations of tests of OPOWER initiatives have not addressed these critical challenges in a systematic way.

Given the complications of experimental design and the magnitude of incentive payments determined by the CPUC, PG&E will engage an expert consultant to design and administer the surveys to provide the highest possible response rate so that energy savings attribution calculations are credible to all stakeholders. The development of the measurement approach that will be used will involve a combination of direct mail surveying of large household samples in combination with in-home surveys designed to observe and correct for any non-response bias that may have occurred. Power analysis will be conducted to ensure adequate sample sizes to detect any differences between groups if they do exist.

Conclusion

Prior evaluations of OPOWER initiatives underscore the potential of home energy reports in stimulating energy savings—and the pitfalls of estimating their unique contribution to behavior change in today's complex environment. PG&E intends to leverage this past work and advance the state-of-the-art in its execution of the largest experiment of the OPOWER reports to date.