Lawrence Berkeley National Laboratory
Recent Work

Title
Insights from Smart Meters: Ramp-up, dependability, and short-term persistence of savings from Home Energy Reports:

Permalink
https://escholarship.org/uc/item/0c56315n

Authors
Todd, Annika
Perry, Michael
Smith, Brian
et al.

Publication Date
2015-04-21

eScholarship.org Powered by the California Digital Library
University of California

Lawrence Berkeley National Laboratory
Environmental Energy Technologies Division **Behavior Analytics**
Providing insights that enable evidence-based, data-driven decisions

Insights from Smart Meters:
Ramp-up, dependability, and short-term persistence of savings from Home Energy Reports

AUTHORS:

Annika Todd[†] Michael Perry, Brian Smith[††], Michael Sullivan[‡], Peter Cappers[†], Charles Goldman[†]

[†]Energy Analysis and Environmental Impacts Division
Lawrence Berkeley National Laboratory

[‡]Nexant

[††]Pacific Gas & Electric Co.

April 2015

behavioranalytics.lbl.gov LBNL-182265

Acknowledgments

The U.S. Department of Energy funded this document to support the State and Local Energy Efficiency Action Network's Customer Information and Behavior (CIB) Working Group. The working group is chaired by Rebecca Wagner, Nevada Public Utilities Commission. The federal staff leads for the Working Group are Michael Li, U.S. Department of Energy and Stacy Angel, U.S. Environmental Protection Agency.

This report was prepared with highly valuable input, direction and comment by members of the CIB Working Group and other technical experts, including: Jim Stewart, Susan Mazur-Stommen, Rebecca Wagner, Lisa Schwartz, Kira Ashby, Aimee Savage, Brian Urban, Abigail Daken, Alex Orfei, Anne Dougherty, Ram Narayanamurthy, Nicholas Payton, Nick Lange, and Richard Caperton.

The work described in this report was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (DOE EERE) under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

Disclaimer

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or simply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California. Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

FOR MORE INFORMATION

On this report:

Michael Li
U.S. Department of Energy
michael.li@ee.doe.gov

Annika Todd
Lawrence Berkeley National Lab
atodd@lbl.gov

On the LBNL Behavior Analytics team:

behavioranalytics.lbl.gov
Annika Todd, Anna Spurlock, Peter Cappers
Lawrence Berkeley National Lab
atodd@lbl.gov, caspurlock@lbl.gov, pacappers@lbl.gov

This document was final as of April 21, 2015. If referenced, it should be cited as:

State and Local Energy Efficiency Action Network (2015). *Insights from Smart Meters: Ramp-Up, Dependability, and Short-Term Persistence of Savings from Home Energy Reports.* Prepared by Todd, A; Perry, M.; Smith, B.; Sullivan, M.; Cappers, P.; and Goldman, C. of Lawrence Berkeley National Laboratory, LBNL-182265, http://emp.lbl.gov/publications/insights-smart-meters-ram

Smart Meter Data: the Opportunity

Smart meters, smart thermostats, and other new technologies provide previously unavailable high-frequency and location-specific energy usage data. Many utilities are now able to capture real-time, customer specific hourly interval usage data for a large proportion of their residential and small commercial customers. These vast, constantly growing streams of rich data (or, "big data") have the potential to provide novel insights into key policy questions about how people make energy decisions.

What can we do with all of these data? The richness and granularity of these data enable many types of creative and cutting-edge analytics. Technically sophisticated and rigorous statistical techniques can be used to pull useful insights out of this high-frequency, human-focused data. In this series, we call this "behavior analytics." This kind of analytics has the potential to provide tremendous value to a wide range of energy programs.

For example, disaggregated and heterogeneous information about actual energy use allows energy efficiency (EE) and/or demand response (DR) program implementers to target specific programs to specific households; enables evaluation, measurement and verification (EM&V) of energy efficiency programs to be performed on a much shorter time horizon than was previously possible; and may provide better insights into the energy and peak hour savings associated with EE and DR programs (e.g., behavior-based (BB) programs).

In this series, "Insights from Smart Meters," we present concrete, illustrative examples of findings from behavior analytics research using these data that are immediately useful and relevant, including:

- **Proof-of-concept analytics techniques** that can be adapted and used by others;
- **Novel discoveries** that answer important policy questions; and
- **Guidelines and protocols** that summarize best practices for analytics and evaluation.

The goal of this series is to enable evidence-based and data-driven decision making by policy makers and industry stakeholders, including program planners, program administrators, utilities, state regulatory agencies, and evaluators. We focus on research findings that are immediately relevant.

Focus on: ramp-up, dependability, and short-term persistence of savings

In this report, we use smart meter data to analyze the ramp-up, dependability, and short-term persistence of savings in one type of behavior-based (BB) program: Home Energy Reports (HERs). In these programs, reports are mailed to households on a monthly, bi-monthly or even quarterly basis. The reports provide energy tips and information about how a household's energy use compares to its neighbors. HERs typically obtain 1% to 3% annual electricity savings, and several studies report that savings from mature HERs persist over multiple years while the programs are running (and decay after the reports are discontinued).[1]

Analytics Technique

Smart meter data enables savings estimates for each day after each report

Questions remain as to the *short-term* persistence of savings. How quickly do HERs ramp-up–how many days until we see savings? How reliable are the savings in the first few months–are there savings every day, and do they decay over time between reports? Currently, there is less information about these questions.[2]

Key Results

Our results show one example of a HER savings resource with:

- **Quick ramp-up**: savings in 2 weeks
- **Dependability**: relatively stable savings every day between reports

Implication: less frequent reports may increase cost-effectiveness

Why does this matter? Because BB programs are focused primarily on reducing electricity consumption through behavioral changes, there is concern that these savings may be less dependable day-to-day than savings from installation of energy efficient equipment. This uncertainty may pose a barrier to broader deployment of BB programs as an energy efficiency and/or demand response resource because system planners and regulators may not see these programs as a dependable

[1] For example, see: Khawaja and Stewart (2014), DNV GL (2014), Todd et al. (2014), Stewart (2013), Integral Analytics (2012), KEMA (2012), Allcott (2011), Allcott and Rogers (2013), Opinion Dynamics (2012).

[2] Allcott and Rogers (2013) report a pattern of "action and backsliding," in which customers start saving energy within days of receiving a report, but then slowly return to their original energy use between reports.

resource. Our analytics technique uses easily available data to determine the ramp-up and dependability of HER program savings over the short-term (day-to-day), which can help utilities, program planners, system planners, regulators and policymakers:

- **Improve HER program design and reduce deployment costs by optimizing report frequency,** *where reports could be sent out less frequently over time with minimal consequence to the achieved savings levels;*
- **Improve short-term demand and overall energy forecasts,** *where daily savings can be predicted with a reasonable degree of accuracy, resulting in more effective hedging strategies for fuel and purchased power procurement;*
- **Improve HER cost-effectiveness,** *as program costs can be reduced and program benefits can be more accurately predicted.*

Analytics Technique

Smart meter data allows us to estimate the savings from HERs on each day after each report was mailed out. Our analytics technique compares the daily electricity use of the treatment group (i.e., those who received the HERs) to the daily electricity use of the control group (i.e., those who did not receive the HERs). We estimate the savings separately for each day after each report was mailed out.[3] This analysis is complex for a few reasons:

- Every household may not receive their reports at the same time. Reports could be sent out based on any number of factors: day of month, address, bill date, etc.; these mailing dates could vary across customers throughout the month. For our test-case program rollout (discussed in more detail below), reports were sent out based on billing dates.
- The number of days between reports need not be constant; each report may be mailed with a different number of days between them. For example, for our test-case program rollout, there were four weeks between the first-to-second, and second-to-third mailings but then 56 days between all subsequent mailings.

Because of these complexities, in order to estimate the savings for each day after the mailing of each report, we cannot necessarily simply estimate the savings during each calendar date. Instead, we need to align the various mailing dates of different customers in order to estimate the savings on the first, second, third, etc. day after each report was mailed, even if those days are associated with different calendar dates.[4] Note that this alignment presents a challenge as to what segment of control group customers is appropriate to use as a comparison group for treatment group customers that receive their reports on a certain day. We solved this issue by estimating "predicted mailing dates" for control customers based on their billing dates; see the Appendix for more information. We estimated savings for every consecutive day, including weekends.

While mailing the reports out at different times to different customers requires a more complex analysis technique than if all reports were sent out on the same date with the same time period between mailings, it does have one advantage–it better controls for variation in impact over time that may be caused by external temporal factors (e.g., savings may increase or decrease as the daylight and weather changes between report mailings; a difference in mailing dates helps

[3] The appendix describes the regression technique and provides summary statistics and validation of randomization.

[4] For example, if one customer were mailed a report on January 1st, and another customer were mailed a report on January 7th, the first day after the report was mailed would be on January 2nd and 8th, respectively

wash out these differences).[5]

We use data from one particular program rollout as a test-case: we draw upon electricity data from the Pacific Gas & Electric (PG&E) smart meter system to analyze the daily impacts of their Home Energy Reports behavior-based program.

The design of this HER program involves mailing reports to households on a monthly or bi-monthly basis. The letters provide information about the household's energy use in addition to how their energy use compares to their neighbors. The letters also include some energy savings tips. These HER programs are designed as randomized controlled trials (RCTs): households are randomly assigned to either the treatment group that receives the reports, or the control group that does not. A well-designed RCT is the "gold standard" of program evaluation design, and thus allows us to produce unbiased estimates of the energy savings each day.[6]

We analyze hourly interval electricity consumption data for one particular HER program pilot rollout (called the "Gamma Wave" by PG&E). It includes 145,000 households, across all electricity usage levels (other rollouts typically target the top 75% of energy users). Households were drawn from five geographic regions in PG&E's service territories. The PG&E Gamma Wave rollout began in November 2011, with reports being delivered at different times to different groups of customers starting in December 2011 and continuing roughly through the next six months.[7]

[5] However, the assignment of customers as to which day during the month they were mailed their report was not random, it was based on the customers' billing dates. This means that reports received at different times during the month do not perfectly control for external temporal factors to the extent that customers with one bill date change their reaction to the treatment over time in a way that is different than the change in reaction over time for customers with another bill date. For example, customers who have a bill date at the beginning or end of the month may be very different than customers with bill dates in the middle of the month. These unobservable differences may be a cause for different response to the receipt of HERs.

[6] In addition to RCTs, there are other factors that are needed to produce valid energy savings estimates; see Todd et. al 2012.

[7] Because participating customers received reports based on bill dates, customers received their first and subsequent reports at different points during the month.

Analytics Technique: Examine short-term savings persistence

Smart meter data allows us to use an analytics technique that estimates the savings from a HER program on each day after each report is mailed out.

Implication: This technique may help program administrators and evaluators understand the ramp-up, dependability, and short-term persistence of savings in BB programs, which can lead to improved program design, increased cost-effectiveness, and better short-term forecasts.

New Results: Insights from the data

We estimate the savings on each day after each report is mailed in order to gain insights into the short-term persistence of savings. For example, we estimate the savings on the first, second, third, etc., day after the first report ("Report 1") is mailed, savings on the first, second, third, etc. day after Report 2 is mailed, and so on.

First, we examine ramp-up: after the very first report is sent, how soon do we see savings? Results for savings estimates on each day after the first report are shown in Figure 1 (along with the 95% confidence intervals in dotted lines).

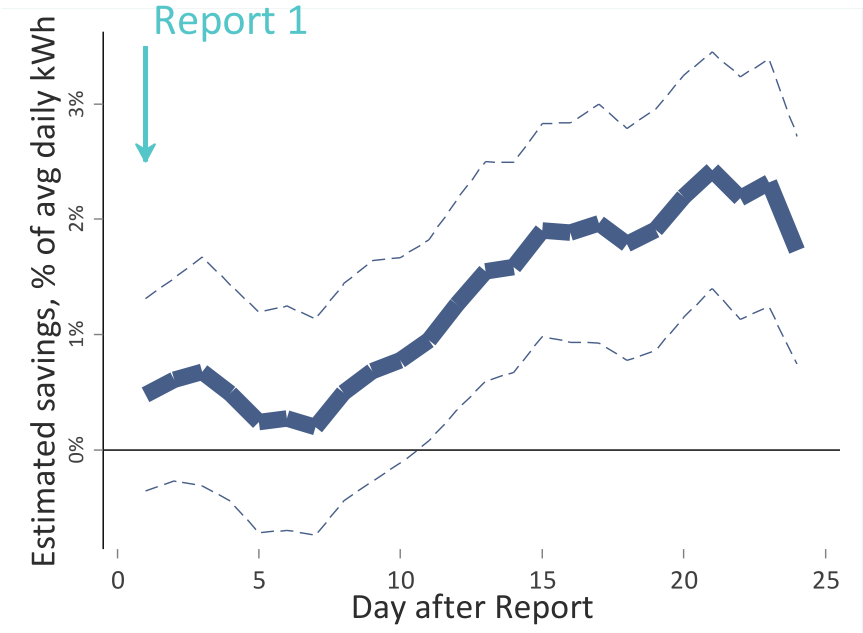

Figure 1. Savings on each day after the first report

The y-axis displays savings as a percent of the average daily energy usage of the control group, the x-axis shows each day after Report 1.

Key Result 1: Quick ramp-up: savings within two weeks

After the first report is mailed out, savings appear to increase rapidly after one week, and are statistically significant after 2 weeks.

Implication: Once deployed, HER programs can be a fast-acting resource for reducing electricity consumption. This is especially true with respect to traditional EE programs such as whole house retrofits, which typically involve a lengthy process of several months between the time when customers get an energy audit, decide on a retrofit package, have contractor that installs and commissions measures, and customer observes energy savings on their utility bill.

Next, we examine the short-term dependability of savings over time: do the savings persist between mailings, and are there savings every day? Do the savings decay or grow over time? Results for savings estimates on each day after the first four reports are shown in Figure 2.

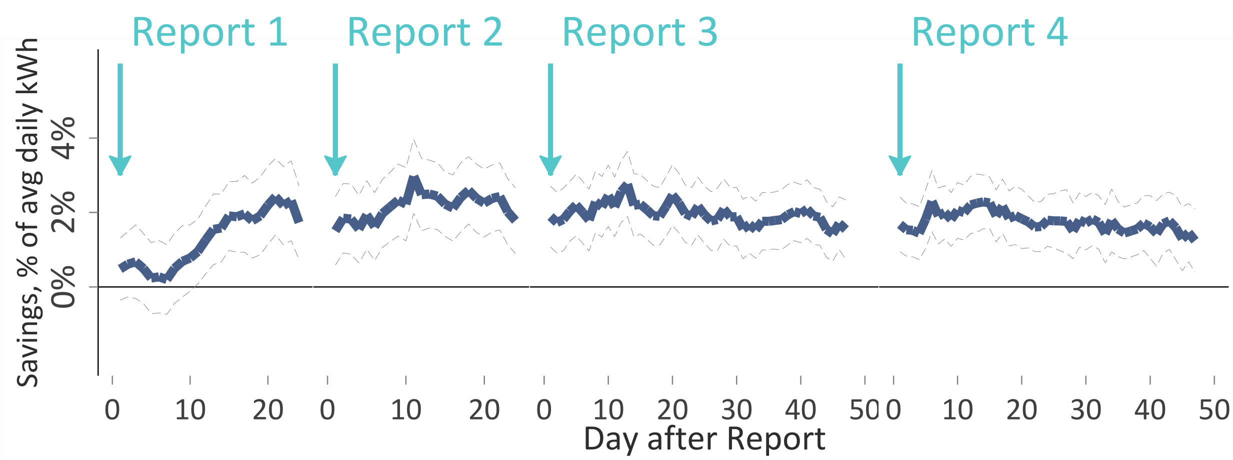

Figure 2. Savings each day after the first, second, third, and fourth report (first 6 months)

The y-axis displays savings as a percent of the average daily energy usage of the control group, the x-axis shows each day after each Report, and dotted lines indicate 95% confidence intervals. Note that although there may appear to be variation across days, and that the savings may appear to be decreasing slowly over time, neither of these effects are statistically significant.[8]

[8] That is not to say that daily variations or slow reductions over time in the savings level do not exist, but rather we simply don't have sufficient power and precision to say that either are occurring.

Key Result 2: **Savings are reliable** – they persist between mailings and are relatively stable

Savings persist between mailings: there are statistically significant savings every day between mailings. The savings are relatively stable: after the first mailing, there is no statistically significant growth or decline in savings over time, and no statistically significant variation in savings day-to-day.

Implications:

- Savings from HERs appear to persist and provide a stable resource for load reduction in the short-term; this is useful information for system and program planning as well as load procurement and forecasting.
- Because the savings appear to stabilize, and do not significantly decline in the eight weeks between mailings 3 and 4, or 4 and 5, it may be possible to increase the duration between reports without affecting the level of savings. This would likely improve cost-effectiveness; this should be tested.

How does this relate to other studies? Allcott and Rogers (2013) estimate daily savings for a HER rollout in the Pacific Northwest; they find a pattern of "action and backsliding," in which customers start saving energy within days of receiving a report, but then slowly return to their original energy use. They note that this is consistent with the idea that the reports "cue" customers to remember to perform day to day energy savings actions, such as turning off the lights when leaving the house. We do not see the same results as Allcott and Rogers: that is, we do not see backsliding back to original energy usage levels between reports. The difference in results between this report and their findings may be a difference in the customer base, the year that the program was rolled out, or other external factors. It may also be because our HER program started on different days of the month for different customers, allowing us to partially control for changes in savings due to daily seasonal affects.

Studies looking at multi-year persistence (e.g., Khawaja & Stewart (2014), DNV GL (2014)) have found that savings increase over the first few years. We look only at short-term persistence, over six months; our results do not speak to whether or not there is an increase in savings over the first few years.

Next Steps & Future Research

In this report we discussed analytic techniques that can be used to provide insights into the ramp-up, dependability, and short-term persistence of HERs. Our results suggest that savings ramp-up quickly and are relatively reliable and stable. Our results may be specific to this particular program in this specific situation. Because we only have data from one utility, with a limited set of time-series data, we do not suggest that these results can be generalized to all HER programs.[9] It is important to use these analytics methods for other iterations of this program type in order to draw broader conclusions.

Future research with more data could examine the ramp-up and dependability of savings between mailings on a longer time horizon, for different HER programs, and for different BB programs more broadly.

One important implication of this research that should be tested with future HER programs is the optimal frequency of report delivery, which may be utility or program specific. Because our research suggests that the savings appear to stabilize and do not significantly decline in the two month gaps between Reports 3, 4, and 5, it is possible that the savings would not decline if there were larger gaps between those reports (as well as later reports). For example, we suggest testing the effectiveness of a program with one month between Reports 1, 2, and 3; three months between 3, 4, and 5; four months between 5 and 6; and larger gaps in between subsequent reports. An increase in the duration between reports as the program progresses may significantly improve the cost-effectiveness of HER programs depending on how the program implementer providing the HER is compensated.[10] This may be true even if there are slightly less savings.

This series will continue to explore the kinds of insights that can be pulled from the newly available data captured by smart meters and other sources, and to report our key findings in this series *Insights from Smart Meters*.

[9] In other words, even though the RCT design ensures that the results are *internally valid* (e.g., unbiased for a particular program, with a given participant population and a given time frame) does not mean that the results are *externally valid* (e.g., can be generalized and applied to new populations, circumstances, and future years).

[10] If there is a cost per report, this might improve the cost-effectiveness. For a different business model in which HER programs are compensated based on savings, this may not improve cost-effectiveness.

References

Allcott, Hunt. 2011. "Social Norms and Energy Conservation." Journal of Public Economics , Vol. 95, No 9-10 (October), pages 1082-1095.

Allcott, Hunt. March 2014. "Site Selection Bias in Program Evaluation." Working paper, R&R in the Quarterly Journal of Economics.

Allcott, Hunt, and Todd Rogers. 2013. "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation." National Bureau of Economic Research Working Paper"

Allcott, Hunt, and R. Sweeney. March 2014. "Information Disclosure Through Agents: Evidence From a Field Experiment." 2014 POWER Conference.

DNV GL. July 2014. "Home Electricity Report Program, January 2012 through December 2013 Study Period: 2013 Impact Evaluation, Seattle City Light."

EPRI. 2011. "The Effect on Electricity Consumption of the Commonwealth Edison Customer Application Program Pilot: Phase 1." EPRI, Palo Alto, CA: 2011. 1022703.

Freeman, Sullivan & Co. 2012. "2011 Load Impact Evaluation for Pacific Gas and Electric Company's SmartAC Program," San Francisco, CA.

Integral Analytics (2012). Impact & Persistence Evaluation Report: Sacramento Municipal Utility District Home Energy Report Program.

Khawaja, Sami M., and Stewart, James. 2014. "Long-Run Savings and Cost-Effectiveness of Home Energy Report Programs." The Cadmus Group, Waltham, MA. CAD10312014.

KEMA. 2010. "Puget Sound Energy's Home Energy Reports Program: 20 Month Impact Evaluation." Madison, WI.

KEMA. 2012. "Puget Sound Energy's Home Energy Reports Program: Three Year Impact, Behavioral and Process Evaluation." Madison, Wisconsin: DNV KEMA Energy and Sustainability.

Opinion Dynamics (2012). "Massachusetts Three Year Cross-Cutting Behavioral Program Integrated Report." Waltham, MA: Opinion Dynamics Corporation.

SMUD. 2013. "SmartPricing Options Interim Evaluation: An interim evaluation of the pilot design, implementation, and evaluation of the Sacramento Municipal Utility District's Consumer

Behavior Study."

Stewart, James. Work in progress, Nov 2013. "Peak-Coincident Demand Savings from Residential Behavior-Based Programs: Evidence from PPL Electric's Opower Program."

Smith, Brian A., and Lucy Arnot, August 2014. "Neighbor Comparison Reports Produce Savings, but HOW?" ACEEE Summer Study 2014. Working Paper.

Todd, A., M. Perry, B. Smith, M. Sullivan, P. Cappers, and C. Goldman. 2014. *Insights from Smart Meters: The Potential for Peak Hour Savings from Behavior-Based Programs.* Lawrence Berkeley National Laboratory, LBNL-6598E. http://emp.lbl.gov/publications/insights-smart-meters-potential-peak-hour-savings-behavior-based-programs.

Todd, A., E. Stuart, S. Schiller, and C. Goldman. Prepared for State and Local Energy Efficiency Action Network. 2012. "Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations." http://behavioranalytics.lbl.gov.

Lawrence Berkeley National Laboratory
Environmental Energy Technologies Division **Behavior Analytics**
Providing insights that enable evidence-based, data-driven decisions

APPENDIX

Insights from Smart Meters: Ramp-up, dependability, and short-term persistence of savings from Home Energy Reports

AUTHORS:

Annika Todd[†] Michael Perry, Brian Smith[††], Michael Sullivan[†], Peter Cappers[†], Charles Goldman[†]

[†]Energy Analysis and Environmental Impacts Division
Lawrence Berkeley National Laboratory

[†]Nexant

[††]Pacific Gas & Electric Co.

April 2015

behavioranalytics.lbl.gov LBNL-182265

Acknowledgments

The U.S. Department of Energy funded this document to support the State and Local Energy Efficiency Action Network's Customer Information and Behavior (CIB) Working Group. The working group is chaired by Rebecca Wagner, Nevada Public Utilities Commission. The federal staff leads for the Working Group are Michael Li, U.S. Department of Energy and Stacy Angel, U.S. Environmental Protection Agency.

This report was prepared with highly valuable input, direction and comment by members of the CIB Working Group and other technical experts, including: Jim Stewart, Susan Mazur-Stommen, Rebecca Wagner, Lisa Schwartz, Kira Ashby, Aimee Savage, Brian Urban, Abigail Daken, Alex Orfei, Anne Dougherty, Ram Narayanamurthy, Nicholas Payton, Nick Lange, and Richard Caperton.

The work described in this report was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (DOE EERE) under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231.

Disclaimer

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or simply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California. Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

FOR MORE INFORMATION

On this report:

Michael Li
U.S. Department of Energy
michael.li@ee.doe.gov

Annika Todd
Lawrence Berkeley National Lab
atodd@lbl.gov

On the LBNL Behavior Analytics team:

behavioranalytics.lbl.gov
Annika Todd, Anna Spurlock, Peter Cappers
Lawrence Berkeley National Lab
atodd@lbl.gov, caspurlock@lbl.gov, pacappers@lbl.gov

This document was final as of April 21, 2015. If referenced, it should be cited as:

Todd, A., M. Perry, B. Smith, M. Sullivan, P. Cappers, and C. Goldman. 2014. *Insights from Smart Meters: Ramp-up, dependability, and short-term persistence of savings from Home Energy Reports.* Lawrence Berkeley National Laboratory, LBNL-182265, http://emp.lbl.gov/publications/insights-smart-meters-ram

Table of Contents

Appendices

These Appendices provide detailed descriptions as an addendum to the paper: "Insights from Smart Meters: Ramp-up, dependability, and short-term persistence of savings from Home Energy Reports." In Appendix A, we provide a detailed description of Home Energy Reports (HERs) and the experimental design (a Randomized Controlled Trial, (RCT)). Appendix B describes the data used in the analysis, and Appendix C provides summary statistics and a validation of the randomization. In Appendix D we describe our analytical approach and present the results in a table format (graphical representations are available in the main body of the paper).

Appendix A: Program description and experimental design

In this section we provide an overview of Opower's Home Energy Reports program that was implemented at PG&E, the program design employed, and a general overview of our analysis methods and the available data.

A.1 Description of Home Energy Reports

Opower worked with PG&E to provide its residential customers with periodic Home Energy Reports (HERs) by mail that contain energy usage feedback and behavioral suggestions (see Figure A-1 for an example). Specifically, the HER compares a customer's monthly electric and/or gas usage to an average of similar homes' usage as well as to an average of the most efficient 20% of similar homes' usage. These "neighbor comparisons" are based on a variety of customer characteristics, including location, home floor area, presence of high energy consuming devices (e.g., pool), and type and number of air conditioning and/or heating units.

The neighbor comparison is used to give the customer one of three ratings:

- *Great* – the customer is more efficient than both average neighbors and efficient neighbors
- *Good* - the customer is more efficient than average neighbors but less efficient than their efficient neighbors
- *Using More than Average* - the customer is less efficient than both average and efficient neighbors

If a customer receives a rating of "Good" or "More than Average," the HER will include a dollar amount of savings that the customer could realize on their annual energy bills by matching their efficient neighbors' usage. A HER also provides a list of several energy savings tips and their potential annual dollar savings. For customers receiving reports on their electric usage, the reports include a graph of their load shape by hour for an average day from the last month of usage. Load shapes are not provided for natural gas usage because gas usage data are generally not collected hourly.

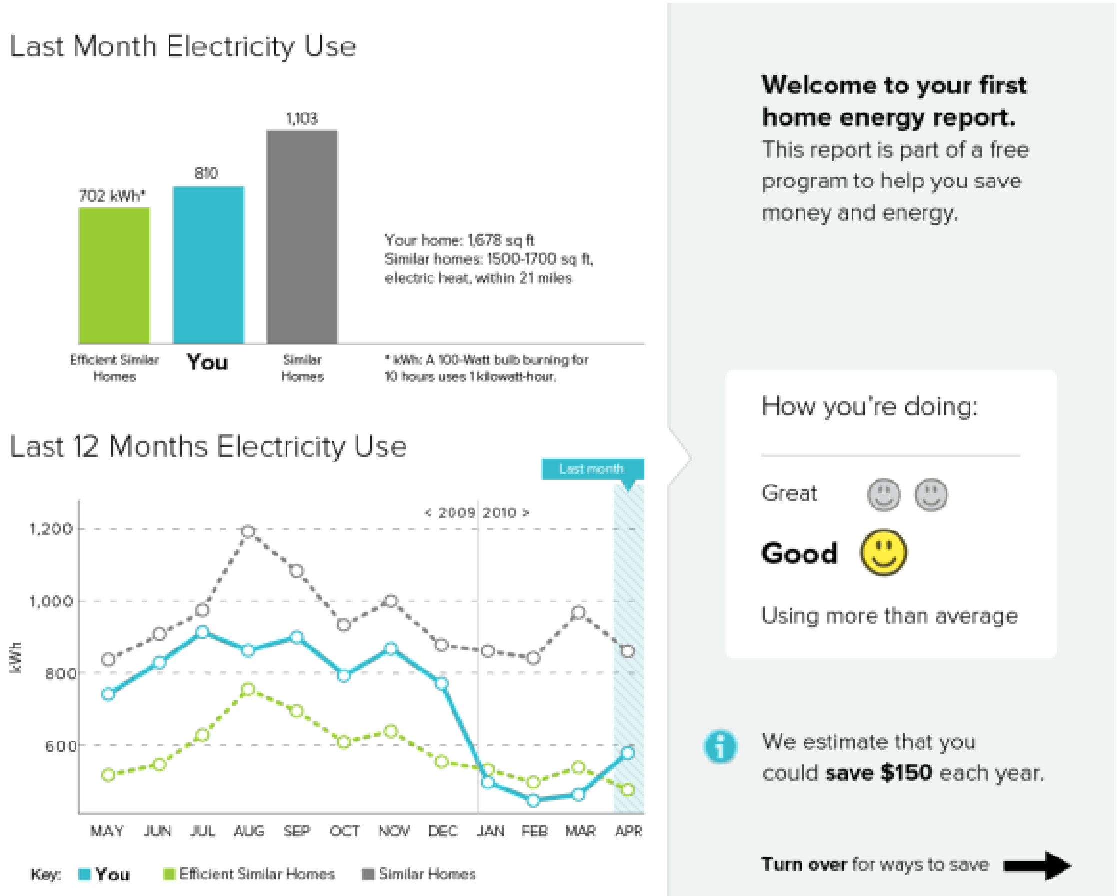

Figure A-1. Example of a Home Energy Report

A.2 Experimental Design

Opower's HER program in PG&E's service territory was designed as a field experiment that employed a randomized controlled trial (RCT). An RCT is a type of experimental design in which households in a given population are randomly assigned to two groups: a treatment group that receives the reports and a control group that does not.

The HER program utilizes an opt-out recruiting process. HERs are sent out to customers assigned to the treatment group without their prior knowledge or approval. These customers can elect to opt-out of receiving future HERs, if they wish by contacting PG&E.[1] Customers in the treatment group can then decide for themselves if and how to best respond to the energy usage feedback and behavioral suggestions contained in the HER. Customers in the control group are likely not aware that an experiment is occurring, since they are likely unaware their

[1] PG&E reports that the HERs generate very few complaints and opt-outs.

peers in the treatment group are receiving HERs, and are therefore unlikely to become dissatisfied.

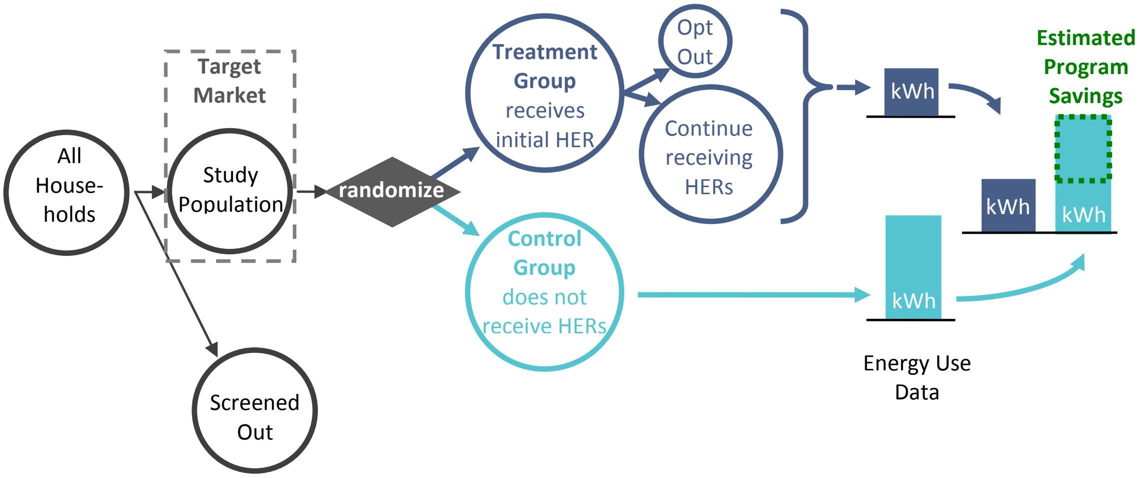

Figure A-2. Experimental design: opt-out randomized controlled trial

Because HERs are designed as RCTs, we can readily compare energy use data from customers in the treatment group to those in the control group in order to produce valid and unbiased statistical estimates of the total electricity savings, the peak demand savings, and the hour-by-hour electricity savings.

A.3 Screening criteria

PG&E's residential customers were screened into the study population using certain inclusion criteria (in addition to satisfying geographic or energy usage criteria discussed in Appendix B:). Customers must: have a full year of bills (to provide pre-treatment data for savings estimation); have had a functioning smart meter for more than one year; be on selected rate schedules—either PG&E's standard residential rate schedule or one of its residential time-of-use rates; neither be on a medical baseline rate, nor flagged as "vulnerable or disabled" in PG&E databases; not be master metered;[2] not be net metered;[3] not live in a mobile home; not be on an electric vehicle rate; not be on a natural gas vehicle rate; not be in another HER pilot program; not live in a multifamily dwelling; not be billed by a municipality; and have not previously requested that PG&E cease sending them any and all marketing materials.

[2] Master metered means that several homes share one meter—such as in a trailer park.

[3] Net metered homes have the ability to generate as well as consume power.

Appendix B: Data description

In this study, we analyze hourly interval electricity consumption data for one particular HER program pilot rollout within the broader set of HER programs implemented in PG&Es service territory (called “Gamma Wave” by PG&E; see Table B-1). It includes around 145,000 households, from all energy usage levels, drawn from five geographic regions in PG&E’s service territories (see Figure B-1 for more information about PG&E’s geographic territories). The Gamma Wave rollout began November 4, 2011, and we obtained data from the beginning of the rollout to September 31, 2012.

Table B-1. Overview of the Wave One dataset

	# Treat	# Control	Launch Date	Hourly interval data available	PG&E territory	Quartile of energy use[4]	Service received from PG&E
Gamma Wave	72,300	72,3000	Nov 2011	Nov 4, 2011 – Sept 31, 2012	R, S, T, W, X	All quartiles	Electric & gas service, and electric-only service

[4] The top (4th or highest) quartile refers to the 25% of energy users who use the most total annual energy on average (using the most energy as compared to the rest of the population). The quartiles were determined based on a combined electric and gas usage index.

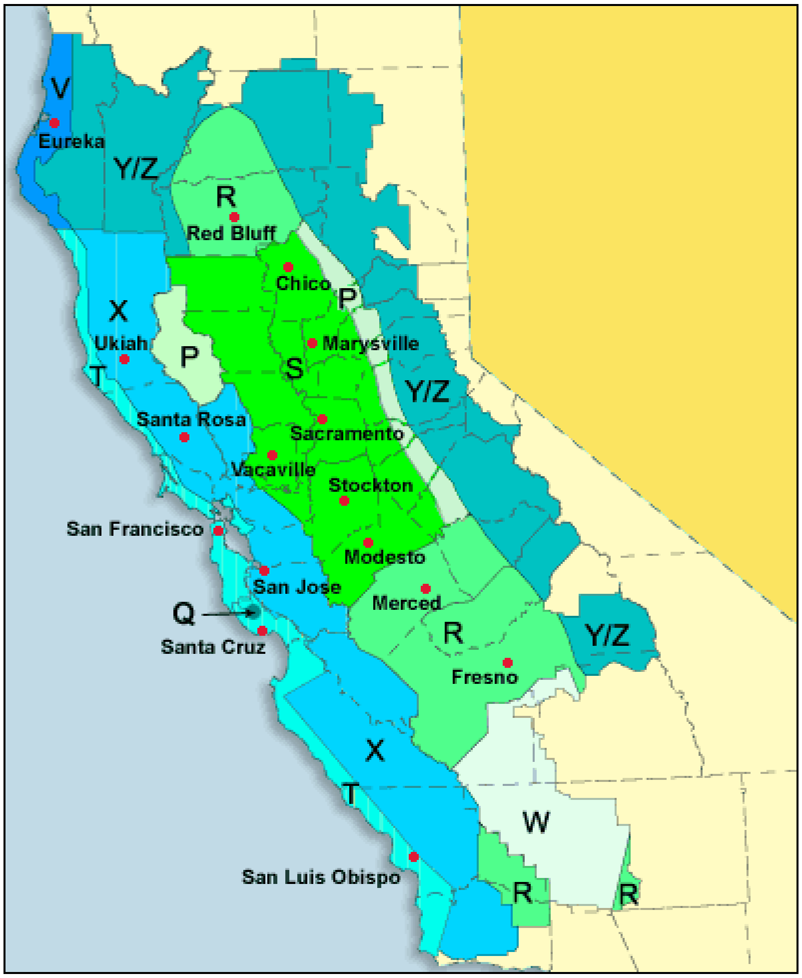

Figure B-1. PG&E Territory Map

Appendix C: Descriptive statistics and validation of randomization

In this section we present descriptive statistics of the pilot and pre-pilot study waves, and validate the comparability between the control and treatment groups through randomization.

Table C-1 demonstrates the successful randomization of customers onto control and treatment groups, as well as showing basic summary statistics. The table shows both the percentage of customers with observed characteristics as well as mean values for quantitative variables.[5] The observed characteristics in the table include baseline territory, CARE status (a program for low-income households offering subsidized rates), income level as estimated by a third party, homeownership status as estimated by a third party, home attributes, and monthly electricity usage prior to treatment. As the table shows, the distribution of each characteristic is similar across treatment and control groups.

The table also shows the results of statistical tests that tell us whether there is any evidence that the distribution of a given characteristic is correlated with treatment status. For binary variables, a z-test on the difference in means was used and the p-value for equality of means is shown. For metrics with more than two categories, the test used was Fisher's exact test and the p-value for independence of category with respect to treatment and control is shown.

Table C-1 shows the number of customers that were sent the first mailing in each wave; the number of months since wave inception through December 2012; and the average monthly attrition rate due to account closure from the beginning of the wave through December 2012. It is our understanding that account closure occurs primarily due to customers moving. In our analysis, we assume that moving (and any other source of account closure) is independent of being in the treatment or control groups.

[5] Data for tables C-1 and C-2 come from a combination of PG&E and third party databases licensed by PG&E.

Table C-1. Distributions of Characteristics across Treatment and Control Groups (Gamma Wave)

<table>
<tr><th>Metric</th><th>Category</th><th>Unit</th><th>Treatment</th><th>Control</th><th>P-value</th></tr>
<tr><td rowspan="5">Baseline Territory</td><td>R</td><td>(% of group)</td><td>22.0%</td><td>22.0%</td><td rowspan="5">1.00</td></tr>
<tr><td>S</td><td>(% of group)</td><td>21.2%</td><td>21.2%</td></tr>
<tr><td>T</td><td>(% of group)</td><td>18.0%</td><td>18.0%</td></tr>
<tr><td>W</td><td>(% of group)</td><td>22.0%</td><td>22.0%</td></tr>
<tr><td>X</td><td>(% of group)</td><td>16.8%</td><td>16.8%</td></tr>
<tr><td colspan="2">CARE Rate</td><td>(% of group)</td><td>36.6%</td><td>36.6%</td><td>0.91</td></tr>
<tr><td rowspan="4">Estimated Household Income</td><td><$30k</td><td>(% of group)</td><td>20.8%</td><td>20.8%</td><td rowspan="4">0.43</td></tr>
<tr><td>$30k-$50k</td><td>(% of group)</td><td>18.1%</td><td>18.2%</td></tr>
<tr><td>$50k-$80k</td><td>(% of group)</td><td>30.1%</td><td>30.5%</td></tr>
<tr><td>>$80k</td><td>(% of group)</td><td>31.0%</td><td>30.6%</td></tr>
<tr><td colspan="2">Renter Status</td><td>(% of group)</td><td>6.8%</td><td>6.8%</td><td>0.91</td></tr>
<tr><td colspan="2">Presence of Pool or Spa</td><td>(% of group)</td><td>13.8%</td><td>13.8%</td><td>0.69</td></tr>
<tr><td colspan="2">Estimated Number of Residents</td><td>(number of residents)</td><td>2.7</td><td>2.7</td><td>0.16</td></tr>
<tr><td colspan="2">Living Space</td><td>(square feet)</td><td>1651.7</td><td>1649.2</td><td>0.71</td></tr>
<tr><td colspan="2">Year Home Built</td><td>(year)</td><td>1968.6</td><td>1968.4</td><td>0.21</td></tr>
<tr><td colspan="2">Estimated Age of Head of Household</td><td>(years)</td><td>53.3</td><td>53.3</td><td>0.95</td></tr>
</table>

Pre-HER Usage	**Oct-10**	**(monthly kWh)**	558	555	0.21
	Nov-10	**(monthly kWh)**	531	529	0.26
	Dec-10	**(monthly kWh)**	597	595	0.31
	Jan-11	**(monthly kWh)**	575	574	0.40
	Feb-11	**(monthly kWh)**	493	492	0.31
	Mar-11	**(monthly kWh)**	518	516	0.20
	Apr-11	**(monthly kWh)**	477	476	0.24
	May-11	**(monthly kWh)**	508	507	0.40
	Jun-11	**(monthly kWh)**	675	673	0.42
	Jul-11	**(monthly kWh)**	834	831	0.45
	Aug-11	**(monthly kWh)**	836	833	0.39
	Sep-11	**(monthly kWh)**	718	716	0.46
	Oct-11	**(monthly kWh)**	558	556	0.29

Table C-2. Monthly Attrition Rate

Wave		Gamma
# of Customers at Launch of Wave	Control	70,529
	Treatment	70,518
# of Months of HERs*		14
Monthly Rate of Attrition (%)	Control	0.9%
	Treatment	0.9%

Appendix D: Analysis and results

In this section, we describe our analytical approach used to estimate the savings on each day after each report is mailed.

D.1 Creating "predicted mailing dates" for the control group

Every household did not receive their reports at the same time; they were mailed out based on billing dates. To estimate the savings for each day after the mailing of each report, we aligned the various mailing dates of different customers in order to estimate the savings on the first, second, third, etc. day after each report was mailed, even if those days are associated with different calendar dates. This alignment presents a challenge as to what segment of control group customers is appropriate to use as a comparison group for treatment group customers that receive their reports on a certain day, because billing dates may not be random within customer segments. We solved this issue by estimating "predicted mailing dates" for control customers based on their billing dates.

Fortunately, we have billing dates for all of the customers. The vast majority (97.6%) of treatment and control group customers fall into one of 15 billing groups, with between 10,000-20,000 customers in each group. For example, all customers in billing group A received their bills on August 4, September 2, October 3, November 5, etc. The 3.4% of customers that did not fit into one of the 15 groups were billed at what appeared to be random dates and were dropped from our dataset.

Within each billing group, almost every treatment customer was mailed reports on the same dates. We define these dates as *predicted mailing dates* for each billing group. Customers in the control group were also given predicted mailing dates according to the billing group that they are in; these are the dates that they would have been mailed reports had they been in the treatment group.

However, some customers were sent reports on dates that did not match the rest of their billing group: 4.65% of customers had actual mailing dates that were different than the predicted mailing dates of their billing group. Because we do not know why these customers were taken out of their billing group and sent reports on different days, and therefore cannot determine which control customers should also be taken out of their billing group, in all analyses we preserve the observable billing group by using the predicted mailing dates rather than actual mailing dates.

D.2 Specification and results

We used the following regression specification, which was estimated separately for each mailing m (so there were four regressions, for $m = 1,2,3,4$): [6]

$$kwh_{id} = \sum_{d=1}^{D} \beta_d^m D_d^m T_i + \sum_{d=1}^{D} \alpha_d^m D_d^m + \varepsilon_{it} \tag{0.1}$$

Where:

- i indicates each household;
- m indicates each mailing;
- d indicates each day after each mailing m;
- kwh_{id} indicates electricity use for household i on day d;
- D_d^m is an indicator variable for each day d after each mailing m;
- T_i is an indicator variable for households in the treatment group;
- β_d^m is the coefficient of interest: the estimated average treatment effect for each day d after each mailing m; and
- Standard errors are robust and clustered at the household level, within each regression, to account for correlation within customers across days.

Table D-1 displays the results; a graphical representation is in the main body of the report.

[6] Pre-treatment data was not available and thus we could not perform a difference-in-differences approach. Because this is a randomized controlled trial, we would expect that adding pre-treatment data for a difference-in-difference analysis would increase the precision but not affect the estimates of savings.

Table D-1. Savings estimates on each day after each report is mailed

	Days after Report 1 mailed kwh		Days after Report 2 mailed kwh		Days after Report 3 mailed kwh		Days after Report 4 mailed kwh	
1 days after	-0.1484*	(.0595)	-0.2910***	(.0657)	-0.3455***	(.0584)	-0.3144***	(.0521)
2 days after	-0.2193**	(.0709)	-0.2986**	(.0941)	-0.2823***	(.0704)	-0.3041***	(.052)
3 days after	-0.2405**	(.0796)	-0.2565*	(.1251)	-0.2896***	(.0697)	-0.2880***	(.0517)
4 days after	-0.1509*	(.067)	-0.3358***	(.0727)	-0.3653***	(.0574)	-0.2805***	(.0509)
5 days after	0.0682	(.136)	-0.3491***	(.0758)	-0.5134***	(.0994)	-0.6437*	(.2572)
6 days after	0.0101	(.1356)	-0.3208***	(.0747)	-0.4646***	(.0952)	-0.5050	(.278)
7 days after	-0.1097	(.0669)	-0.3570***	(.0653)	-0.3323***	(.0591)	-0.3458***	(.0537)
8 days after	-0.1505*	(.0673)	-0.3818***	(.0665)	-0.4000***	(.0587)	-0.3583***	(.0509)
9 days after	-0.2087**	(.0743)	-0.4127***	(.0878)	-0.3819***	(.0714)	-0.3293***	(.0512)
10 days after	-0.2391***	(.0653)	-0.3271**	(.1187)	-0.4312***	(.0718)	-0.3612***	(.0531)
11 days after	-0.2123***	(.0623)	-0.5030***	(.0823)	-0.3819***	(.0571)	-0.3579***	(.0527)
12 days after	-0.1002	(.1385)	-0.4403***	(.0819)	-0.5227***	(.101)	-0.4165	(.2292)
13 days after	-0.1913	(.1398)	-0.4508***	(.0752)	-0.5602***	(.0985)	-0.4254	(.2637)
14 days after	-0.3016***	(.0649)	-0.4212***	(.0641)	-0.3861***	(.0606)	-0.3962***	(.0514)
15 days after	-0.3462***	(.0659)	-0.3936***	(.0617)	-0.3913***	(.0571)	-0.3953***	(.0492)
16 days after	-0.3769***	(.0755)	-0.3566***	(.0896)	-0.3663***	(.0683)	-0.3287***	(.0502)
17 days after	-0.3913***	(.0841)	-0.3650**	(.1289)	-0.3124***	(.0667)	-0.3705***	(.0515)
18 days after	-0.3282***	(.0717)	-0.4490***	(.0729)	-0.3455***	(.0547)	-0.3376***	(.0528)
19 days after	-0.2655	(.1431)	-0.4470***	(.074)	-0.4497***	(.0932)	-0.3704	(.2553)
20 days after	-0.2361	(.1429)	-0.4435***	(.0724)	-0.4849***	(.0891)	-0.3893	(.2343)
21 days after	-0.4202***	(.0732)	-0.4144***	(.064)	-0.3959***	(.0571)	-0.3228***	(.0498)
22 days after	-0.3859***	(.0749)	-0.4219***	(.0642)	-0.3561***	(.055)	-0.3043***	(.0472)
23 days after	-0.4293***	(.0815)	-0.3416***	(.0951)	-0.3228***	(.0683)	-0.3149***	(.0488)
24 days after	-0.3688***	(.0733)	-0.3394*	(.1366)	-0.3351***	(.0671)	-0.3223***	(.0507)
25 days after					-0.3526***	(.0545)	-0.3270***	(.0521)
26 days after					-0.4194***	(.0947)	-0.4040	(.2424)
27 days after					-0.4445***	(.095)	-0.7441**	(.2698)
28 days after					-0.3737***	(.0578)	-0.2896***	(.0532)
29 days after					-0.3434***	(.056)	-0.3309***	(.0521)
30 days after					-0.3191***	(.0691)	-0.2867***	(.0486)
31 days after					-0.2819***	(.0693)	-0.3014***	(.0503)
32 days after					-0.3149***	(.055)	-0.3219***	(.0514)
33 days after					-0.3931***	(.0905)	-0.2581	(.2204)
34 days after					-0.3849***	(.0855)	-0.3182	(.2217)
35 days after					-0.3277***	(.0542)	-0.2987***	(.053)
36 days after					-0.3312***	(.0548)	-0.2810***	(.0495)
37 days after					-0.2949***	(.0734)	-0.2754***	(.0508)
38 days after					-0.3238***	(.0712)	-0.2714***	(.0495)
39 days after					-0.3644***	(.056)	-0.3170***	(.0525)

40 days after			-0.5006***	(.0971)	-0.9212*	(.3595)
41 days after			-0.4751***	(.0977)	-0.6955*	(.343)
42 days after			-0.3480***	(.0542)	-0.3107***	(.0566)
43 days after			-0.3434***	(.0537)	-0.3244***	(.0549)
44 days after			-0.2691***	(.0663)	-0.2675***	(.0552)
45 days after			-0.2600***	(.0628)	-0.2171***	(.0527)
46 days after			-0.3178***	(.0508)	-0.2818***	(.0561)
47 days after			-0.4052***	(.0886)	-0.6583*	(.328)
48 days after			-0.3775***	(.0909)	-0.5293	(.3607)
49 days after			-0.3577***	(.0569)	-0.2977***	(.0573)
50 days after			-0.3265***	(.0533)	-0.2680***	(.0536)
51 days after			-0.3322***	(.0656)	-0.3051***	(.0578)
52 days after			-0.3355***	(.0646)	-0.2602***	(.0555)
53 days after			-0.3520***	(.0531)	-0.2582***	(.052)
Day after Mailing FE	Yes	Yes	Yes		Yes	
R-squared	.0091704	.0017687	.003325		.0080698	
Number of Households	193064	193038	193040		192995	

Standard errors in parentheses
Note: SE clustered at household level for each regression
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix E: References

Allcott, Hunt. 2011. “Social Norms and Energy Conservation.” Journal of Public Economics , Vol. 95, No 9-10 (October), pages 1082-1095.

Allcott, Hunt. March 2014. “Site Selection Bias in Program Evaluation.” Working paper, R&R in the Quarterly Journal of Economics.

Allcott, Hunt, and Todd Rogers. 2013. “The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation.” National Bureau of Economic Research Working Paper”

Allcott, Hunt, and R. Sweeney. March 2014. “Information Disclosure Through Agents: Evidence From a Field Experiment." 2014 POWER Conference.

DNV GL. July 2014. “Home Electricity Report Program, January 2012 through December 2013 Study Period: 2013 Impact Evaluation, Seattle City Light.”

EPRI. 2011. “The Effect on Electricity Consumption of the Commonwealth Edison Customer Application Program Pilot: Phase 1.” EPRI, Palo Alto, CA: 2011. 1022703.

Freeman, Sullivan & Co. 2012. “2011 Load Impact Evaluation for Pacific Gas and Electric Company’s SmartAC Program,” San Francisco, CA.

Integral Analytics (2012). Impact & Persistence Evaluation Report: Sacramento Municipal Utility District Home Energy Report Program.

Khawaja, Sami M., and Stewart, James. 2014. “Long-Run Savings and Cost-Effectiveness of Home Energy Report Programs.” The Cadmus Group, Waltham, MA. CAD10312014.

KEMA. 2010. “Puget Sound Energy’s Home Energy Reports Program: 20 Month Impact Evaluation.” Madison, WI.

KEMA. 2012. "Puget Sound Energy’s Home Energy Reports Program: Three Year Impact, Behavioral and Process Evaluation." Madison, Wisconsin: DNV KEMA Energy and Sustainability.

Opinion Dynamics (2012). "Massachusetts Three Year Cross-Cutting Behavioral Program Integrated Report." Waltham, MA: Opinion Dynamics Corporation.

SMUD. 2013. “SmartPricing Options Interim Evaluation: An interim evaluation of the pilot design, implementation, and evaluation of the Sacramento Municipal Utility District’s Consumer

Behavior Study."

Stewart, James. Work in progress, Nov 2013. "Peak-Coincident Demand Savings from Residential Behavior-Based Programs: Evidence from PPL Electric's Opower Program."

Smith, Brian A., and Lucy Arnot, August 2014. "Neighbor Comparison Reports Produce Savings, but HOW?" ACEEE Summer Study 2014. Working Paper.

Todd, A., M. Perry, B. Smith, M. Sullivan, P. Cappers, and C. Goldman. 2014. *Insights from Smart Meters: The Potential for Peak Hour Savings from Behavior-Based Programs.* Lawrence Berkeley National Laboratory, LBNL-6598E. http://emp.lbl.gov/publications/insights-smart-meters-potential-peak-hour-savings-behavior-based-programs.

Todd, A., E. Stuart, S. Schiller, and C. Goldman. Prepared for State and Local Energy Efficiency Action Network. 2012. "Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations." http://behavioranalytics.lbl.gov.