Do LED lightbulbs save natural gas? Interpreting simultaneous cross-energy program impacts using electricity and natural gas billing data

Oluwatobi G Adekanye1,∗, Alex Davis1 and Inês L Azevedo1,2

1 Department of Engineering & Public Policy, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, United States of America
2 Department of Energy Resources Engineering, Stanford University, 367 Panama St, Stanford, California, 94305, United States of America

∗ Author to whom any correspondence should be addressed.
E-mail: oga@alumni.cmu.edu

Keywords: energy efficiency, buildings, event history modeling, difference-in-differences modeling, fixed-effects modeling

Abstract

Local and state governments have made significant advances in creating, implementing, and evaluating energy efficiency programs in the buildings sector. Evaluations commonly use ex-post statistical models to complement ex-ante engineering estimates when determining program impacts. A critical assumption of data-driven evaluations is that reductions would not occur in the absence of the program. This assumption is difficult to test, particularly if other unobserved changes to the building’s energy profile coincide with the program’s adoption. We provide a method to detect a class of unobserved simultaneous cross-energy changes in a building’s energy profile by examining the treatment effects of electricity-only programs on natural gas use and vice versa. We apply the method to a panel of residential energy efficiency implementations with monthly electricity and gas data from 2010 to 2016 in the City of Palo Alto, California. Using difference-in-differences and event history analyses, we find evidence of significant gas reductions estimated for some electricity-only programs, suggesting that households implemented unobserved changes at the same time as those programs. Our results highlight how data-driven analyses may not adequately estimate program impacts, and the value of simultaneous electricity and natural gas measurements for detecting and interpreting unobserved changes to energy use at the household level. Lastly, we present evidence that energy savings from non-monetary interventions can exceed those which offer financial rewards for energy efficiency.

Introduction

Energy efficiency is a cost-effective way to reduce energy use, with billions of dollars invested annually in the private and public sectors [1–3]. Because buildings are responsible for about 41% of total US energy consumption and around one-third of CO₂ emissions, there is the potential for using building energy efficiency as a tool for long-term energy and carbon reduction goals [4–15]. While local and state governments across the US have implemented many programs to promote building energy efficiency, there is always a question of whether programs meet their targets.

Two important methods are frequently used to quantify the energy savings of energy efficient technologies, each with its strengths and weaknesses. The first is deemed savings: for example, the savings of switching from an incandescent to a light emitting diode (LED) lightbulb can be estimated by multiplying the difference in energy use per hour between the incandescent and the LED bulb by the number of hours the incandescent bulb was used. However, this analysis is only accurate if use remains the same. Behavioral responses such as the rebound effect (where LEDs are more frequently used than CFLs), as well as implementation challenges (where energy efficient appliances are not installed correctly), would violate that assumption, and in turn, invalidate the
We divide the dataset for our analysis into 3 parts: exploratory (20%), training (60%), and test (20%) to avoid problems with overfitting. We use 20% of the dataset for exploratory analysis to generate models that fit the patterns in the data as closely as possible. We preregistered the models using the Open Science Framework (https://osf.io/jtnqf/), then used cross-validation to test these models against each other on the training 60% of the data. We choose the model with the lowest cross-validation error to predict the last 20% of the data (test set).
Statistical inference for the candidate models (standard errors, confidence intervals) is calculated using the last 80% of the data. We focus on 80% of the dataset (excluding exploratory data) for the rest of this paper.

Figure 1 shows the average electricity consumption (in kWh/day) and gas consumption (in therms/day) for our sample in the City of Palo Alto from January 2010 to December 2016.

Table 1. Energy Efficiency program description in the CPA between 2010 and 2016 [36].

| Energy efficiency program               | Description                                                                                                                                 |
|----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| Smart Energy                           | provides rebates to residents who install energy efficient appliances and equipment in their homes or on their property. These include home heating and cooling systems, insulation, water heaters, pool pumps, and power strips |
| Residential Energy Assistance Program (REAP) | provides home lighting and heating system upgrades (such as insulation for walls and roofs and weather stripping for doors and windows) to low-income residents at no cost. |
| Refrigerator Recycling                 | discourages the use of second refrigerators by providing rebates to customers who sustainably recycle their old refrigerators through the city’s program. |
| Green @Home Acterra                    | offers free in-home audits through a local nonprofit volunteer environmental organization (Acterra). At the end of the audit, participants receive personalized efficiency tips along with simple efficiency-improvement items such as compact fluorescent lamps (CFLs), faucet aerators, and home energy monitors (for larger consumption customers). |
| CFL Bulb                               | provides residents with free globe style CFLs at local stores. The CPAU mailed coupons to residents that were presented at the store in exchange for CFL bulbs. |
| LED 2/$8                               | provides rebates for up to two LED bulbs from $38 each to $4 where coupons were mailed to residents. |
| LED Holiday Light                      | exchanges a working strand of incandescent holiday lights for a free LED strand through the use of coupons which residents will present to local hardware stores. |
| Home Energy Kit                        | provides a home energy efficiency kit to customers. Includes energy saving appliances such as new CFL bulb, high energy efficiency nightlight, and water saving faucet aerators. |

Statistical inference for the candidate models (standard errors, confidence intervals) is calculated using the last 80% of the data. We focus on 80% of the dataset (excluding exploratory data) for the rest of this paper.

Figure 1 shows the average electricity consumption (in kWh/day) and gas consumption (in therms/day) for the middle 60% of the data (approximately 16,000 households). In the CPA, we find aggregate coincident peaks of gas and electric use in the coldest month of the year, January, where homes use the most heating. Almost 60% of homes in California use natural gas for heating while another 21% use electricity, which would cause both electricity and natural gas use to increase during the winter [37]. The difference in usage rates for natural gas versus electricity is mirrored in figure 1, which shows a large percentage difference between seasons for natural gas than for electricity. We use a log-transformation for both electricity and gas consumption as there are heavy right tails in their distributions. Figure 2 shows the average electricity and gas consumption information for the year 2016 with and without the log transformation respectively where the log transformation yields a more normal distribution pattern. The patterns of heavy right tails in the base case and normal distribution in the log transformation case are also seen for all years (see the supplementary information section S1 (available online at stacks.iop.org/ERC/3/015003/mmedia) for more details).

From the sample of 16,000 households, 3480 households applied for at least one energy efficiency program between 2008 and 2015. Figure 3 shows the distribution of the different energy efficiency program applications over the timeframe of our analysis and the quarters over which these energy efficiency programs were received. The highest number of energy efficiency programs were received in the second quarter of 2010 when the LED
A $2/S8 program was implemented. However, this program was implemented in 2010 and ended in 2011. Similar patterns are seen for the Home Energy Kit and CFL programs which were implemented over a one to two-year timeframe. This has significant implications for our analysis, as the lack of enough pre-treatment observations may make some estimates unreliable. Other programs like the Smart Energy and LED Holiday lights program are relatively stable over the timeframe of our analysis. Collectively these two programs make up the largest

**Figure 2.** (A) Histogram of average electricity consumption per day in 2016. (B) Histogram of logged average electricity consumption per day in 2016. (C) Histogram of average gas consumption per day in 2016. (D) Histogram of logged average gas consumption per day in 2016.

**Figure 3.** Energy efficiency programs implemented by the City of Palo Alto between 2010 and 2015.
percentage of energy efficiency programs received (about 86% of the total program enrollment) over the timeframe of consideration. Overall, the appliance rebate programs (i.e. Smart Energy program) and lighting programs (LED Holiday Light, LED 2/$8) account for the highest percentage of energy efficiency programs received. Refrigerator Recycling accounts for a much smaller percentage ~8% of total energy efficiency programs while Green@home Acterra (i.e. education of residents on green at home practices) accounts for only approximately 6% of total energy efficiency programs received (The SI section S1 contain more details).

Modeling strategy
The decision to participate in an energy efficiency program is voluntary, which raises concerns about issues such as selection bias, which occurs when program participants share characteristics that makes them different from non-participants, leading to bias in study estimates. Participating households, for example, may be more concerned about environmental issues, where a study in 2017 found that more than six in ten adults favor California making its own policies to address global warming [38]. This will lead to bias if those households that participated in the program also have a different trajectory of energy use before the implementation of the energy efficiency program compared to those who did not adopt the program.

To account for these issues, we use two models: (1) difference-in-differences, and (2) event history, following earlier studies, including Ito, Fowlie et al Boomhower and Davis, Novan and Smith, and Zivin and Novan [18, 19, 32, 39, 40]. First, we implement the difference-in-differences model by comparing the change in electricity and gas consumption for households that received an energy efficiency program to those that did not receive an energy efficiency program, adjusting for time-varying covariates (such as weather). Using fixed effects, we examine these changes to reduce the effect of time-invariant unobserved differences between households that may affect both energy efficiency program applications and electricity or gas usage. We start with the difference-in-differences model as:

\[
\ln (\text{kWh}_{pt}) = \alpha_e + \beta_e (\text{House}_p) + \gamma_e (\text{MonthNumber}_t) + \sum_r \theta_e r (\text{EE}_pt) r + \varepsilon_{e, pt}
\]

\[
\ln (\text{therms}_{pt}) = \alpha_g + \beta_g (\text{House}_p) + \gamma_g (\text{MonthNumber}_t) + \sum_r \theta_g r (\text{EE}_pt) r + \varepsilon_{g, pt}
\]

In equations (1) and (2), kWh and therms represent the electricity and gas consumption per day of household p in month t where t ranges from month 1 to month 84 (as we have data over a 7-year time frame). House_p and MonthNumber_t are vector dummy variables which represent the household and month-in-timeframe fixed effects, with \( \beta \) and \( \gamma \) represented as a vector of these dummy variables with subscripts for both electricity and gas respectively. As weather patterns are observed monthly do not vary by region as all data are obtained from the same area, we capture weather patterns by including the MonthNumber_t variable. (EE_{pt})_r is an indicator variable for the different energy efficiency programs, \( r \), as described in table 1 above. These are 0/1 indicator variables accounting for pre and post energy efficiency implementation trends. They take on a value of 1 after the rebate program has been issued for household p in time period t, and 0 otherwise. \( \theta_e, \) and \( \theta_g \) are our coefficients of interest. They represent the changes in electricity and gas use respectively over all the energy efficiency programs, \( r \), as described in table 1 above. \( \varepsilon_{e, pt} \) and \( \varepsilon_{g, pt} \) represents the error term for both electricity and gas consumption respectively.

A key assumption when implementing the difference-in-differences model is that the treatment groups have similar trends to the control groups in the absence of the treatment. In our case, we assume that those who receive an energy efficiency program would follow a similar trend to the control group if they had not received the program—known as the parallel trends assumption. The difference-in-differences model also assumes that the electricity and gas reductions roughly follow a step function (evidenced by the 0/1 pre and post treatment indicators).

Because we have seven-years worth of data, we can examine the pre and post-treatment trends to ensure that the treatment and control groups do not violate the parallel trends assumption—which we refer to as the event history analysis. As opposed to differences-in-differences, which has the 0/1 pre and post-treatment indicator, the event history analysis specifies monthly indicators to denote time windows around the months the rebate programs are issued. Therefore, we can examine individual monthly trends to see if the treatment and control groups follow similar trends before the rebates are issued and if these trends continue or change after implementation. Also, as mentioned in the data section above, some energy efficiency programs could have been implemented months in advance of rebate program issuance. The event history analysis also allows us to examine if this is indeed the case, as the evidence of significant dips before the rebate issuance date does not only violate the parallel trends assumption but also may indicate the evidence of treatment months in advance of rebate issuance.

As different households receive the energy efficiency program at different times, we implement the event history analysis as follows: we standardize the modeling framework to account for time 0 as the month a
household receives an energy efficiency program as well as specify a 1–12 month indicator window before and after energy efficiency program implementation for the different energy programs. Months after the 12-month window are coded as 12+ while months before the 12-month window coded as −12+. We model the event history with equations (3) and (4) as follows:

$$\ln (\text{kWh}_{pt}) = \alpha_\epsilon + \beta_\epsilon (\text{House}_p) + \gamma_\epsilon (\text{MonthNumber}_t) + \sum_r \left( \sum_{n=-12+}^{1} \varphi_{r,e} (K_{pn}) \right) + \varepsilon_{e,pt}$$  \hspace{1cm} (3)

$$\ln (\text{therms}_{pt}) = \alpha_g + \beta_g (\text{House}_p) + \gamma_g (\text{MonthNumber}_t) + \sum_r \left( \sum_{n=-12+}^{1} \varphi_{r,g} (K_{pn}) \right) + \varepsilon_{g,pt}$$  \hspace{1cm} (4)

In equation (3) and (4), kWh and therms represent the electricity and gas consumption per day of household p in month t where t ranges from month 1 to month 84 (as we have data over a 7-year time frame). House, and MonthNumber, represent the vectors of household and month-in-time-frame fixed effects to account for time-invariant household and month effects with β and γ represented as a vector of these dummy variables for both electricity and gas respectively. As the different rebate programs are issued at different times in the timeframe, j specifies monthly indicator windows around rebate program issuance where j could take on values ranging from −12+, −12, −11…, −3, −2, −1. T_{j} are interactions of the energy efficiency program indicator, which equals 1 if household p ever adopted the specific energy efficiency program and time dummies for all periods before month 0. Likewise, K_{pn} is the treatment indicator interacted with time dummies for all time periods after month 0 with n taking on values of 1, 2, 3…11, 12, 12+. ε_{e,pt} and ε_{g,pt} represents the error term for both electricity and gas consumption respectively.

For all models in this study, we cluster the standard errors at the household level to account for autocorrelation between errors of the households over different months.

Robustness checks
We conduct several robustness checks to ensure the accuracy of our program estimates using difference-in-differences, comparing models (1) and (2) in the previous section with the results of all our robustness checks. We use 3 checks: (a) seasonality effects, (b) ‘long’ and ‘short’ run effects, and (c) a shuffle test. Firstly, we examine the sensitivity of the model estimates to account for seasonality effects as energy efficiency program impacts may be stronger in the winter, for example, compared to the summer. We subset our data into tertiles of low, medium, and high temperature days (corresponding to the heat degree days in the CPA) for electricity and gas consumption. We then re-run Models (1) and (2) using the smaller subsets to examine its effects. Second, we estimate ‘long’ and ‘short-run’ effects to ensure that the effects captured are the result of the energy efficiency program and no other factors. We term ‘long-run’ effects as effects captured too far into the future that may be as a result of other factors and not the implementation of the energy efficiency program. For example, an energy efficiency program may be implemented leading to little or no reduction in energy consumption, but a major home remodel may occur in the same household months in the future resulting in significant energy reductions and hence, overall significant average reductions in energy consumption and misattribution. We term ‘short-run’ effects as immediate reductions that might be seen as a result of the energy efficiency program implementation. To implement this, we subset households with at least one full year of pre and post treatment data while we subset households with at least 2 full years of data for those that never got a program. We then re-run (1) and (2) to examine the long-run effects. In the short-run case, we re-run models (1) and (2) but for only the 12-month window. The SI section S2 provides more details.

We also use a ‘shuffle test’ where we redistribute information on control and treatment households to inspect if we get non-significance of our model estimates. The shuffle test mixes the treatment and control group information such that households that originally receive a treatment become controls (i.e. behave as if they did not receive a treatment) while some control households receive a treatment using a random selection process. As the number of households in the control group is much higher than that in the treated group, we randomly shuffle the control group multiple times so that different households which originally were not treated randomly ‘receive’ a treatment in different iterations. We expect that this mixing would yield no significant reductions in electricity or gas use if the energy efficiency program indeed has an effect. However, if we see reductions with this mixing method, then it indicates a problem with the model. The SI section S3 provides additional details.
Due to publicly available reports under the SB 1037 bill implemented by the State of California, we also have access to the annual reports documented by the CPA for specific energy efficiency programs. As a result, we can compare the savings from our data-driven models with savings from the evaluation, measurement, and verification process (EM&V) presented by the CPA for some energy programs available during the timeframe of our analysis. Finally, we compare the difference-in-differences and event history model using the remaining 20% of the data. We make predictions about the test data comparing both models and calculate the residuals to get the root-mean square error (rMSE) (See the SI sections S4 and S5 for more details and results).

**Limitations**

While we have attempted to accurately capture the effects of the different energy efficiency programs using different model specifications as well as employing varied robustness checks, our model is subject to some major data limitations. Without access to customer or other demographic information of the households, it is difficult to test hypotheses about the cause of cross-energy effects. Secondly, the availability of higher-resolution data at the 15-min or even hourly intervals could have allowed for the use of more sophisticated tools to tease out individual appliance(s) and hence program effects. Finally, more targeted questions surrounding the rebate programs, such as appliance installation date, could have allowed for better accuracy when estimating program impacts.
Results

Energy savings of energy efficiency programs
Figures 4(a), (b) shows the regression results on electricity and gas consumption respectively from Models (1) and (2) in the Methods section (see the SI section S2 for full model results).

From figures 4(a), (b), we find varied effectiveness of the energy efficiency programs. Of all the programs, the Green@Home Acterra program is associated with the largest reductions in both electricity and gas usage at $-6\%$ (95%CI: $-11\%$ to $-2\%$) and $-6\%$ (95%CI: $-10\%$ to $-2\%$) respectively. We find significant reductions for electricity but not gas usage at $-2\%$ (95%CI: $-3\%$ to $-1\%$) and 1% (95%CI: $-1\%$ to 2%) respectively for the Smart Energy program while the REAP low-income program does not show significant reductions in either electricity ($-6\%, 95\%$CI: $-14\%$ to 2%) or gas ($-4\%, 95\%$CI: $-11\%$ to 2%) use. From the program description, a lot of appliances qualify for the Smart Energy rebate program and a significant number of appliances can be upgraded in households that qualify for the REAP program indicating that savings can vary widely from household to household. As a result, the wide ranges estimated for the REAP and Smart Energy program may be as a result of the grouping of a large number of appliances which have varied energy savings. Surprisingly, we find increases rather than decreases for the Home Energy Kit program at 4% (95%CI: 0.3% to 7%) for electricity and 4% (95%CI: 0.1% to 8%) for gas, respectively. As the Home Energy Kit program was implemented and ended in 2010, the first year of the time frame for our dataset, there may be a lack of pre-treatment observations to appropriately estimate its effect. We also find that the Refrigerator Recycling program is associated with significant reductions in electricity usage, but also significant increases in gas use, potentially indicating simultaneous unobserved factors (such as installing gas-consuming appliances while replacing the refrigerator). We implement robustness checks in the section below to further examine why this might be the case.
We find cross-energy effects of some electricity-only programs on natural gas usage. The LED 2/$8 program shows no significant reductions for electricity use at $-0.4\%$ (95%CI: $-3\%$ to $2\%$) but a small reduction effect with gas usage at $-3\%$ (95%CI: $-5\%$ to $-0.4\%$). The LED Holiday light program shows significant electricity reductions of $-4\%$ (95%CI: $-6\%$ to $-1\%$) and gas reductions of $-2\%$ (95%CI: $-5\%$ to 0.2%) respectively. Our results suggest that households that receive the LED lighting programs may also be implementing other unobserved changes in their households that could affect both electricity and natural gas usage.

Exploratory investigation of unexpected effects
From the previous section, we find cross-energy effects of the LED 2/$8$ and LED Holiday Light programs on natural gas usage. We hypothesize two reasons that might be artifacts of our method: (1) differences in the number of observations for each month over the timeframe of observation and (2) long and short-run effects. The observed electricity and gas data for most households in our dataset contain missing data for some months. Although these electricity and gas reductions are averaged when estimating different energy efficiency program impacts, there is the possibility that these reductions are unusually higher in some months compared to others.

Secondly, as stated in the methods section, we also hypothesize that significant effects may be estimated as a result of other effects too far into the future that may be captured and not as a result of the implementation of the energy efficiency programs—termed as ‘long-run’ effects. We also compare this effect with the ‘short-run’ effects observed following the implementation of an energy efficiency program. To test our hypotheses, we first subset our dataset to households who have at least 12 months out of 18 months of pre and post-treatment information. We implement this method of subsetting to ensure that we are capturing households with at least one full year of data before and after energy efficiency program implementation. We then rerun Models (1) and (2) to examine the impacts of the different energy efficiency programs in the long-run case. For the short run case, we only examine the effects using the 12-month window pre-and post program implementation to capture immediate reductions as a result of the energy efficiency program implementation. In figures 5(a), (b) we compare results of the base case i.e. using the full dataset (and as evidenced in figures 4(a), (b)), with the long and short-run effects of those with a full year of pre and post-treatment information.

Because the number of households who qualify for these programs reduces as a result of the constraints on the long and short term effects, we examine not only the statistical significance of the programs but also the variation in estimates. From figures 5(a), (b), we find that the majority of the program estimates do not significantly vary as a function of the model specification—the Green@Home Acterra program, for example, varies by about $-1\%$ for electricity and $-3\%$ for gas (For electricity reductions - base: $-6\%$ (95% CI: $-11\%$ to $-2\%$), long-run: $-6\%$ (95% CI: $-10\%$ to $-1\%$), short-run: $-7\%$ (95% CI: $-11\%$ to $-3\%$); for gas reductions - base: $-6\%$ (95% CI: $-10\%$ to $-2\%$), long-run: $-9\%$ (95% CI: $-14\%$ to $-4\%$), short-run: $-7\%$ (95% CI: $-11\%$ to $-2\%$). Concerns about increases in consumption for the Home Energy Kit and Refrigerator Recycling program, however, reduces significantly when examining the long-run and short-run effects, indicating that the large number of post-treatment data and outliers may indeed impact program estimates. The Home Energy Kit electricity estimates, for example, reduces from a 4% (95%CI: 0.3% to 07%) increase in the base case to a $-0.4\%$ (95%CI: $-4\%$ to 4%) and 1% (95%CI: $-2\%$ to 4%) non-significant decrease in the long and short-run estimates respectively for electricity. Gas estimates for the Home Energy Kit program also show equally similar results of a 4% (95%CI 0.1% to 8%) slightly significant increase to a 1% (95%CI: $-4\%$ to 6%) and 1% (95%CI: $-4\%$ to 5%) non-significant decrease in long and short run estimates respectively. The Refrigerator Recycling program gas estimates also reduces significantly from a 6% (95%CI: 2% to 10%) increase in the base case to a 3% (95%CI: $-2\%$ to $-8\%$) and 2% (95%CI: $-2\%$ to 5%) non-significant increase in long and short-run estimates. 95% confidence intervals clustered at the household level.

The LED 2/$8$ program is highly sensitive to the choice of model specification. From figure 5(b), we find that the short-run estimates are significantly smaller than the base and long-run estimates. The LED 2/$8$ program gas estimates go from a $-3\%$ (95%CI: $-5\%$ to $-0.4\%$) and $-3\%$ (95%CI: $-6\%$ to $-0.04\%$) reduction in the base and long-run estimates to a $-1\%$ (95%CI: $-4\%$ to 1%) non-significant reduction in short term estimates. Just like the Home Energy Kit program, the LED 2/$8$ program was implemented in 2010 and 2011, therefore, we hypothesize that we may be capturing other effects in the long-term which may be associated with significant gas reductions. The LED Holiday light program, however, shows more consistent estimates, with reductions in the short term. We find a $-2\%$ (95%CI: $-5\%$ to 0.2%) and $-5\%$ (95%CI: $-7\%$ to $-2\%$) reduction in the LED holiday light gas estimates in the base and long-run case while the short term also shows a $-2\%$ (95%CI: $-5\%$ to 0.2%) reduction.

Event history model
As explained in the methods section, we implement the alternative event history model specification to examine individual monthly trends before and after program implementation. We find that generally, the event history
models for each of the energy efficiency programs roughly approximate a step function indicating that the difference-in-differences may be appropriate for our analyses (the SI section S2 shows the event history plots for the different programs).

**Other robustness checks**

As explained in the methods section, we also explore issues of seasonality as well as conduct a shuffle test to ensure that our model specification is accurate.

**Seasonality effects**

We explore seasonality effects as we are concerned about stronger effects during the summer, compared to the winter, for example. We find that our estimates do not vary significantly as a result of different temperature ranges (see the SI section S2 for more detailed results). The results are somewhat expected, as Palo Alto has fairly temperate weather patterns all year-round. From these results, we are less concerned about issues of seasonality.

**Shuffle test**

We randomly shuffle control households that get a treatment and vice versa, repeat models (1) and (2) regressions 25 times, then average the estimates and robust standard errors. We find that none of the results are significant (see the SI section S3 for more detailed results).

**Comparison with savings from the CPAU**

We compare the electricity and gas savings claimed by some of these energy efficiency programs using annual independent report information in accordance with the State’s 2021 Assembly bill. We could not make comparisons for the Green@Home Acterra and Home Energy Kit program because these savings were not available in the annual reports. We also did not compare savings for the Smart Energy and REAP programs because a variety of appliances potentially qualify for rebates and are eligible for low-income upgrades respectively. Therefore, these savings would vary widely from appliance to appliance. Table 2 compares our modeling estimates with the information available for the remaining programs.

As can be seen from table 2, Refrigerator Recycling estimates are much lower when using the data-driven than the EM&V. One reason for this discrepancy may have been that CPA customers had refrigerators replaced rather than removed, as was assumed in the EM&V. A second reason could be that the old refrigerators used less energy than assumed in the EM&V analyses (based on age, size, features, and usage patterns). Next, CFL Flood Lights and LED 2/$8 savings for our study are well within the range of the EM&V estimates, particularly considering the large uncertainty ranges in the data analysis. We hypothesize that the main reason for this uncertainty is due to the timing of program implementation—both programs were implemented at the beginning of our study period, in 2010 and ended shortly after, in 2011 allowing for a short pre-treatment period with a long post-treatment period. As expected, the LED Holiday Lights estimates are much larger using the data-driven than EM&V. If this is due to spillover effects caused by the LED Holiday Lights program, then the EM&V estimates show a strong downward bias because they fail to take into account the program’s effect on other energy efficiency purchases. On the other hand, if this difference is due to incidental changes to the household appliances or behavior that happen to coincide with the LED Holiday Lights program (but are not caused by the program), then the data-driven LED Holiday Lights data-driven estimate is a large overestimation. The cross–energy effects on natural gas support the conclusion that there is some simultaneous change occurring in households that adopted the program, but cannot differentiate whether those effects are due to spillover caused by the program or are incidental.
Discussion

In this study, we find evidence of simultaneous unobserved changes to a household’s energy profile using cross-energy treatment effects, where we observe significant gas reductions for some electricity-only programs. While ex-ante engineering estimates are being used to estimate the savings from the implementation of new technologies, academic research has begun to use data to examine the predicted versus actual savings, with a majority realizing that actual savings are significantly lower compared to estimates [18, 32, 41–43]. Although varied statistical methods are used to examine the actual impacts of these programs, we find that data-driven analyses can also result in detectably biased estimates. Earlier research has mostly used electricity savings (in kWh) or energy (KBtu) to examine program impacts, as it is expected, for example, that an electricity-only program should ideally only show significant kWh reductions. As randomized controlled trials are difficult to design for energy efficiency programs, it is very difficult to tease out if reductions are indeed a function of program implementation of other factors [44–46].

There are many factors that may induce both electricity and gas reductions in tandem with receiving an electricity-only program. For example, a household may claim the LED lights rebate, when undergoing a major home remodel. Other factors like solar PV installations or the departure of high energy use household members may occur at the same time as households adopt the program. Cross-energy effects could indicate an underestimate of program impacts if they reflect true spillover effects, where a program encourages a household to engage in deeper energy use reductions beyond the program’s initial scope. For example, a program’s marketing efforts could indirectly improve participation in other programs by raising awareness about energy efficiency. Advertising from the LED Holiday lights program may have changed motivation or raised awareness about additional energy efficiency investments (including technologies that reduce natural gas use), encouraging additional energy efficiency investments in the household. On the other hand, cross-energy effects might indicate overestimates of program impacts if households adopted the program coincidentally with other unobserved changes to their energy profile, or if those unobserved changes caused households to adopt the program (reverse causality). Without further surveys, the exact reason for these cross-energy reductions is unknown, but by using natural gas measurements, we show how to provide a first-level indication of their presence. We were able to examine these cross-energy effects because we have access to both electricity and gas use data, but other proxies such as water use, can also be used in situations where gas use is not available.

Our results also contribute to the existing literature on the importance of non-monetary programs in energy efficiency interventions. We find the highest reductions in electricity and gas usage from the Green@Home Acterra program which performs a walk-in energy audit and trains residents on energy consumption reduction methods. Here, we find average reductions of 6% on the average of those who engaged in the energy audit program offered by the CPA. Our results are in line with earlier research that has found reductions in electricity consumption through the use of non-monetary approaches [32, 47–49]. Although it is impossible to detect by using energy consumption data, if households simply changed their behavioral patterns or replaced old equipment to newer more-efficient equipment, we add to the body of existing literature which highlights the importance of information provision versus financial incentives to reducing energy use.

Our results are subject to some limitations which yield opportunities for future research. While we have attempted to disentangle the effects of the program using quasi-experimental methods, we are only able to hypothesize about the causes of cross-energy effects, but are not able to experimentally determine their causes. A comprehensive survey using customer demographic and behavioral information in tandem with a data-driven method can be used to characterize the reasons for opting into a program. To disentangle program spillover effects from coincidental effects, surveys might ask program participants whether they set out to make their home more energy efficient when they came across the program, or whether they came across the program and it inspired them to make their home more energy efficient. One problem with such a retrospective survey is that respondents would need to have significant insight into their motivations, as well as a good recall of those motivations for events that may have happened many years prior.

Energy efficiency programs have the potential to significantly reduce electricity and gas use in buildings. However, data-driven evaluations need to detect potential bias from unobserved changes that occur at the same time as program adoption. Our work, in addition to corroborating existing research on energy efficiency program effectiveness, provides a simple way to detecting those unobserved changes using cross-energy treatment effects.

Acknowledgments

Funding from this work comes from the Climate and Energy Decision Making Center (CEDM) www.cedmcenter.org which is funded through a cooperative agreement from the National Science Foundation
(SES-00949710), as well as from the Wilton Scott Institute for Energy Innovation at Carnegie Mellon University. We would like to thank Bruce Lesch, Shiva Swaminathan, and Christine Tam from the City of Palo Alto Utilities for the provision of access to the data as well as responding to inquiries about the different energy efficiency programs. We would also like to thank Pedro Ferreira, Gabrielle Wong-Parodi, Brian Sergi, and Evan Sherwin for helpful comments during the preparation of this paper.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Oluwatobi G Adekanye https://orcid.org/0000-0003-0027-4346

Inês L Azevedo https://orcid.org/0000-0002-4755-8656

References

[1] 2018 Energy Efficiency, 2018-Analysis and Outlooks to 2040 International Energy Agency
[2] ADVANCED ENERGY NOW 2016 MARKET REPORT: Global and US Markets by Revenue 2011–2015 and Key Trends in Advanced Energy Growth, 2016 (Washington, DC: Navigant Research)
[3] American Council for an Energy Efficiency Economy 2017 Energy Efficiency - Jobs and Investments (https://aceee.org/sites/default/files/ee-jobs-money-031119.pdf)
[4] US EIA 2018 Consumption & Efficiency (Washington, DC: US Energy Information Administration)
[5] US EIA 2018 US Energy-Related Carbon Dioxide Emissions (Washington DC: US Energy Information Administration) 2017
[6] Lima Azevedo I, Morgan M G, Palmer K and Lave L B 2013 Reducing US residential energy use and CO2 emissions: how much, how soon, and at what cost? Environ. Sci. Technol. 47 2302–11
[7] State and Local Energy Efficiency Action Network 2016 Energy Efficiency as a Least-Cost Strategy to Reduce Greenhouse Gases and Air Pollution and Meet Energy Needs in the Power Sector Prepared by: Lisa Schneidt, Greg Leventis, Steven R. Schiller, and Emily Martin Fadlhor of Lawrence Berkeley National Laboratory, with assistance by John Shenot, Ken Colburn and Chris James of the Regulatory Assistance Project and Johanna Zetterberg and Molly Roy of U.S. Department of Energy (https://www7.eere.energy.gov/seaction/system/files/documents/pathways-guide-states-final0415.pdf)
[8] Nadel S, Shipley A M and Elliott R N 2004 The technical, economic, and achievable potential for energy efficiency in the United States: a meta-analysis of recent studies 2004 Proc. of the 2004 American Council for an Energy-Efficient Economy Conf. in Buildings
[9] National Academy of Sciences 2010 Real Prospects for Energy Efficiency in the United States (Washington, DC: The National Academies Press)
[10] Czerny Jon et al 2007 Reducing US Greenhouse Gas Emissions: How Much at What cost? McKinsey & Company
[11] The White House and 2016 United States Mid-Century Strategy for Deep Decarbonization
[12] Granade GHannana Choi et al 2009 Unlocking Energy Efficiency in the U.S. Economy McKinsey & Company
[13] Wilson E, Christensen C, Horowitz S, Robertson J and Maguire I 2017 Energy Efficiency Potential in the US Single-Family Housing Stock Energy Efficiency Potential in the US Single-Family Housing Stock
[14] Gowrishankar V and Levin A 2017 America’s Clean Energy Frontier: the pathway to a safer climate future National Resources Defense Council
[15] Wei M, Nelson J H, Greenblatt J B, Mileva A, Johnston J, Ting M, Yang C, Jones C, McMahon J E and Kammern D M 2013 Deep carbon reductions in California require electrification and integration across economic sectors Environ. Res. Lett. 8 14038
[16] First E 1986 Actual energy savings after retrofit electrically heated homes in the Pacific Northwest Energy 11 299–308
[17] Davis L W, Fuchs A and Gertrud P 2014 Cash for coolers: evaluating a large-scale appliance replacement program in Mexico Am. Econ. J. Econ. Policy 6 207–38
[18] Fowlie M, Greenstone M and Wolfram C 2018 Do energy efficiency investments deliver? Evidence from the weatherization assistance program ‘Q’ J. Econ. 133 1397–644
[19] Io K 2015 Asymmetric incentives in subsidies: evidence from a large-scale electricity rebate program Am. Econ. J. Econ. Policy 7 209–37
[20] Alberini A and Towe C 2015 Information v. energy efficiency incentives: evidence from residential electricity consumption in Maryland Energy Econ. 50 530–40
[21] Allcott H and Greenstone M 2017 Measuring the Welfare effects of Residential Energy Efficiency Programs, National Bureau of Economic Research Working Paper Series, No. 23386 (https://doi.org/10.3386/w23386)
[22] Liang J, Qiu Y, James T, Ruddehl B L, Dalrymple M, Earl S and Castelazo A 2018 Do energy retrofits work? Evidence from commercial and residential buildings in Phoenix J. Environ. Econ. Manage. 92 726–43
[23] Thomas B A, Hausfather Z and Azevedo I L 2014 Comparing the magnitude of simulated residential rebound effects from electric end-use efficiency across the US Environ. Res. Lett. 9 74010
[24] Thomas B A and Azevedo I L 2013 Estimating direct and indirect rebound effects for US households with input–output analysis: I. Simulation Econ. 86 188–98
[25] Thomas B A and Azevedo I L 2013 Estimating direct and indirect rebound effects for US households with input–output analysis: II. Theoretical framework Econ. 86 199–210
[26] Salerno V and Rabbeni G 2018 An extreme learning machine approach to effective energy disaggregation Electronics MDPI AG 7 235
[27] Shin C, Rho S, Lee H and Rhee W 2019 Data requirements for applying machine learning to energy disaggregation Energies MDPI AG 12 1696
[28] Liang M, Meng Y, Lu N, Lubs grandson and Kling A 2019 HVAC load disaggregation using low-resolution smart meter data 2019 IEEE Power and Energy Society Innovative Smart Grid Technologies Conf., ISGT 2019 (Institute of Electrical and Electronics Engineers Inc.)
[29] Stankovic L, Stankovic V, Liao J and Wilson C 2016 Measuring the energy intensity of domestic activities from smart meter data Applied Energy Elsevier Ltd 183 1565–80
[30] Grimes A, Preval N, Young C, Arnold R, Denne T, Howden-Chapman P and Telfar-Barnard I 2016 Does retrofitted insulation reduce household energy use? Theory and practice Source: The Energy Journal 37 165–86
[31] Scheer J, Clancy M and Ni Högán S 2012 Quantification of energy savings from Ireland’s Home Energy Saving scheme: an ex post billing analysis Energy Efficiency 6 35–48
[32] Zivin J G and Novan K 2016 Upgrading efficiency and behavior: electricity savings from residential weatherization programs The Energy Journal 4 37 1–24
[33] Min J, Azevedo I L and Hakkarainen P 2015 Assessing regional differences in lighting heat replacement effects in residential buildings across the United States Applied Energy 141 12–8
[34] Hopkins A S, Lekov A, Lutz J, Rosenquist G and Gu L 2011 Simulating a nationally representative housing sample using energyplus (https://escholarship.org/uc/item/96p9d8dw)
[35] 2014 DEER (Database for Energy Efficient Resources) Master Documentationhttps://cpuc.ca.gov/General.aspx?id=4132
[36] Abendschein A N 2017 Informational Report on the City of Palo Alto Utilities Demand Side Management Annual Report for Fiscal Year 2016 City of Palo Alto Utilities
[37] Energy Information Administration Household Energy Use in California: A Closer Look at Residential Energy Consumption. (https://eia.gov/consumption/residential/reports/2009/state_briefs/pdf/CA.pdf)
[38] Baldassare Mark, Bonner David, Kordus David and Lopes Lunna 2017 Californians & the Environment Public Policy Institute of California
[39] Boomhower Judson and Davis Lucas 2020 Do Energy Efficiency Investments Deliver at the Right Time 12 115–39
[40] Novan K and Smith A 2018 The incentive to overinvest in energy efficiency: evidence from hourly smart-meter data J. Assoc. Environ. Resour. Econ. 5 577–605
[41] Dubin J A, Miedema A K and Chandran R V 1986 Price effects of energy-efficient technologies: a study of residential demand for heating and cooling Rand J. of Econ. 17 310–25
[42] Metcalf G E and Hassett K A 1999 Measuring the energy savings from home improvement investments: evidence from monthly billing data Review of Economics and Statistics 81 516–28
[43] Sebold F D and Fox E W 1985 Realized savings from residential conservation activity Source: The Energy Journal 6 73–88
[44] Boomhower J and Davis L W 2014 A credible approach for measuring inframarginal participation in energy efficiency programs + Journal of Public Economics 113 67–79
[45] Gillingham K and Palmer K 2014 Bridging the Energy Efficiency Gap Insights for Policy from Economic Theory and Empirical Analysis Review of Environmental Economics and Policy 8 18–38
[46] Gerarden T D, Newell R G and Stavins R N 2017 Assessing the energy-efficiency gap Journal of Economic Literature 55 1486–525
[47] Allcott H 2011 Social norms and energy conservation Journal of Public Economics 95 1082–95
[48] Allcott H and Rogers T 2014 The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation American Economic Review 104 3003–37
[49] Considine T J and Sapci O 2016 The effectiveness of home energy audits: a case study of Jackson, Wyoming Resource and Energy Economics 44 2–70