2014

The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load Control Program

Corey Lang
University of Rhode Island, clang@uri.edu

Edson Okwelum
University of Rhode Island

The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load Control Program

Corey Lang[*] and Edson Okwelum
Department of Environmental and Natural Resource Economics
University of Rhode Island

December 3, 2014

Abstract
Demand response is an important tool for utilities to manage load during peak periods. While the effects of demand response programs on peak load reductions are well studied and intuitive, assessments typically fail to recognize the potential for off-peak behavioral responses that may mitigate the total benefits of the program. Using smart meter consumption data on residential air conditioning units enrolled in a direct load control program, this paper examines changes in consumption prior to and after curtailment events. The results suggest substantial increases in off-peak consumption, which reduce energy, monetary, and environmental benefits of the program by over 40%.

Keywords: *Demand Response, Pre-cooling, Snapback, Strategic Behavior, Breakpoints*

1 Introduction

Many electricity utilities are challenged to meet demand during peak consumption, usually during summer afternoons.[1] Further, peak demand has been increasing at an average annual growth rate of 0.8% over the last 10 years (NERC 2013), heightening the need for solutions. Importing capacity from other areas and firing peak generators are two costly solutions; these can increase the cost of generation by $29.2-80/MWh. (EIA 2013). Alternatively, demand response (DR) programs offer the possibility of no or low cost reductions in peak load, and utilities are increasingly implementing these programs. In general, DR seeks to reduce peak load by changing consumption prices or offering incentives to consumers. One type of DR is direct load control (DLC), in which the utility can control usage of appliances for a few hours per day during critical event days.

Impact evaluations have demonstrated large benefits of DR programs (e.g., Herter 2007, Herter et al. 2007, Faruiqui and George 2002, Crew et al. 1995, Matsukawa 2001, Binswanger 2001, Berkhout et al. 2000, Gillingham et al. 2013, Thomas and Azevedo 2013), and comparison of load reduction during event hours show that DLC programs are more effective than programs with time-varying price structures (Wolak 2011, Newsham and Bowker 2010). However, most of these assessments tend to only examine benefits in terms of peak reductions, thereby ignoring behavioral responses and potentially overestimating the total benefits.

The purpose of this paper is to rigorously estimate both peak and off peak changes in consumption to better understand net benefits of DLC programs with no price incentive or behavioral aspect. We examine Pacific Gas and Electric's (PG&E) SmartAC program, which is designed to reduce peak cooling load by directly controlling air conditioning (AC) units of participants.

We use data from a stratified random sample of 294 participating AC units. Because there is no data for a control group, we estimate unit-specific non-linear consumption models and then compare load on event days to predicted load. Importantly, we use data from the summer 2007 training period. In contrast to subsequent summers and normal DR designs, during the training period curtailment days were called for many different temperature levels, not just the hottest days. This aspect means we are not predicting out of sample and we can have greater

---

[1] California ISO defines peak period as being from 7am-10pm, however the peak period we refer to in this paper is narrowly defined to be in line with PG&E's Summer Peak periods which is from 12pm-8pm.

confidence in our econometric evidence. Our methodology is validated by a falsification test in which we find no changes in consumption during peak or off-peak times on non-curtailment days that match the temperature and timing profile of actual curtailment days.

Several key results emerge from our analysis. First, we confirm that the SmartAC program, like other DR and DLC programs, reduces peak load during event days. In this case, peak consumption was reduced 21% on average. However, we also find a 14% increase in consumption in the hours preceding and the hours following an event. Essentially, some load is being displaced from peak to off-peak times. In addition to cutting the energy savings, these behavioral changes mitigate the monetary and environmental benefits of the program by over 40%.

The main contribution of this paper is to show DR energy policies can lead to unintended consequences. We contribute to the emerging body of work on the perverse incentives in energy policies (e.g., Fowlie 2009, Davis and Kahn 2010, Goulder and Stavins 2011). Particularly relevant to this paper are works by Holladay et al. (2014), and Jessoe and Rapson (2014). Holladay et al. (2014) examine changes in energy production when media outlets relay emergency calls from utilities for electricity reduction during off-peak hours on curtailment days. They estimate that behavioral responses lead to an increase in generation cost of about $43.70/MWh to $61.70/MWh in the late morning and early evening on emergency days. We estimate a smaller increase in generation cost, around $21/MWh, but the results largely corroborate each other. In contrast to Jessoe and Rapson (2014) who find positive spillovers of load reduction during peak hours on non-event days due to learning, we find negative spillovers in non-event hours on curtailment days.

2 Background on the SmartAC Program

PG&E's SmartAC program is a DLC program designed to reduce participating households' cooling loads during peak load times. The program works by directly controlling AC during peak hours on designated curtailment days. Curtailment events can only occur between the hours of 11am to 7pm from May 1 to October 31 and only for a total of six hours per day and for no more than 100 hours per cooling season. Before and after each event, the customers have complete control over their AC units.

Typically, curtailment days are called when system load is expected to be burdensome, usually the hottest days of the year. The California Independent System Operator (ISO) may declare a system emergency when an electric-resource generation facility reaches or exceeds a certain heat rate (usually heat rates of 15,000 British thermal units per kW). However, PG&E can also call events to test the devices and for other discretionary reasons such as transmission or distribution system overload. The Appendix gives a more detailed description of the process of initiating a control event.

In contrast to the normal objective of the program, summer 2007 was a training period, and curtailment days were called on days with a variety of temperatures, not just the hottest days of the year. The inclusion of typical summer temperature days is important because it allows the utility to test the operational performance of the control devices.

The SmartAC Program uses paging signals to reduce cooling load of enrolled AC units during times of peak system demand. The control devices are either programmable communicating thermostats (PCTs) and/or DLC adaptive switches (switches). Both technologies receive signals through a paging device but differ by how they lower AC load. Unlike the adaptive switch control devices, the thermostats provide additional functionality to participants. Demand reduction during curtailment events are achieved by either adjusting thermostat temperature set points or limiting the duty cycles of switches in the units. The PCTs implement load reduction by increasing the cooling set point temperature on the thermostat controls when an event signal is received. Increasing the PCT set point ensures that temperature increases are equitably distributed across the population irrespective of house, occupant temperature preference and air conditioner attributes. The switches, on the other hand, reduce AC load by directly controlling how the unit's compressor operates.[2] The SmartAC program was structured such that 70% of the program participants had control switches and 30% smart thermostats. Customers with either device can only opt out of an event by either calling a toll-free number or accessing the program website.[3]

---

[2] Switches generally control air conditioner load by limiting the compressor run-time or duty cycle to a maximum over a period. When activated on curtailment days, the switches use an adaptive algorithm to reduce AC load to a percentage of the load on "learning days". The "learning days" are days with similar characteristics as potential curtailment days and are chosen by the program administrator. The observed duty cycle on these learning days provides an estimate of expected duty cycle on a curtailment day.

[3] While participants had the option to opt out of the program, this option was almost never used by the households (KEMA 2008).

There were two schedules by which load was reduced for those households with PCTs, and PG&E alternated between them. The first schedule set thermostats back one degree every other hour, resulting in a maximum setback of three degrees in each six-hour curtailment event. The second schedule set thermostats back one degree every hour for the first four hours of the event and then held constant. This last strategy was done in an attempt to evenly spread the load reduction resulting in a gradual temperature decline compared to the steep temperature increase from the first strategy. For units with adaptive switches, load reduction during curtailment events was achieved by limiting compressor duty-cycle to a maximum amount in the curtailment period.

Using information from focus groups and similar program evaluation surveys in other locations, participants were recruited into the SmartAC program through direct mail campaigns. Participants in the SmartAC program were first enlisted in the spring of 2007 in San Joaquin County. However, this has been expanded to include all service areas covered by the utility. At the beginning of July 2007 when the first curtailment event was called, the program had enrolled over 2,857 households. The 297 households sampled for this evaluation are located in the city of Stockton and its surrounding areas (San Joaquin County). And as of August 2007, which was the system peak, there were 8,193 participants in the program, which was almost composed exclusively of residential customers with less than 1% commercial customers. The program has subsequently grown in popularity to such an extent that by January 2008, there were 26,000 households enrolled in the program with an additional 22,000 waiting to be supplied control technology. All program participants received a one-off incentive payment of $25 in return for up to 100 hours/year of load control. In addition, participants in the program evaluation group were given a maximum payment of $110 based on the number of questionnaires answered in a survey conducted at the completion of the program's first year. Also, participants in the PCT group were given the thermostats free, with a market value of $200 (KEMA 2008). The goal of the program is to enroll a sufficient number of participants to achieve load reduction of 305 MW with over 95% coming from residential accounts.

3 Methodology

We develop a methodology that estimates AC load usage as a function of weather control (temperature and humidity) and time of day, parameterizing each unit separately. Then we use these predictions to compare actual usage on curtailment days to expected usage. While there are

several methods for estimating baseline consumption patterns and load comparisons, we use a regression-based baseline modeling approach. The wide scale adoption of smart grid meters and availability of high-resolution, hourly or 15-minute energy consumption data has contributed a great deal to improvements in regression-based baseline models (Carrie Armel et al. 2013, Newsham et al. 2011, Santin and Itard 2010, Santin et al. 2009). Mathieu et al. (2011) find that the regression-based baseline model performs better than most models used in evaluating DR programs. In addition, the regression-based model allows for customer-specific response estimates to be used separately for decision making or easily averaged to any degree of aggregation. By estimating parameters for each unit, we better capture unique occupant cooling preferences and differences in physical properties of residences.

We model each AC unit's baseline electric load as a function of time of the day, cooling degree days, relative humidity, heat index and a 631 index. Cooling degree days enters into the model as a 2-breakpoint demand function, illustrated in Figure 1. The first breakpoint is the unit's base or reference temperature, which marks the temperature above which AC is used. The second breakpoint represents the temperature point above which the AC unit runs at 100% duty-cycle. This specification allows for three different regimes of sensitivity to temperature changes. To capture regular patterns of usage for each unit, we use 15-minute interval fixed effects. The heat index is a function of relative humidity and temperature and gives an idea of how hot it is given the relative humidity and actual air temperature. We use an equation that reproduces NOAA's heat index chart to estimate the heat index for all participating units. The 631 index is a weighted, moving average of temperature and is estimated as 60% of a given day, 30% of prior day and 10% of two days prior average temperature readings (California ISO 2013).

Specifically, we estimate the baseline model using only non-curtailment days for each unit as follows:

$$L_{ithd} = \alpha_{it} + \gamma_i B_{ihd} + \delta_i C_{ihd} + \eta_i R_{ih} + \theta_i X_{ih} + \varsigma_i Y_{id} + \varepsilon_{ithd} \tag{1}$$

where L_{ithd} is the observed load in kWh of unit i in 15-minute interval t in hour h of day d, α_{it} is a unit-specific 15-minute interval fixed effect, and B_{ihd} and C_{ihd} are the degrees Fahrenheit above the first and second breakpoints, respectively.[4] R_{ih} is the relative humidity, X_{ih} is the heat index for unit i at hour h and Y_{id} is the 631 temperature index. The breakpoints are defined as follows:

---

[4] We distinguish between 15-minute interval and hour because some variables are only available at the hour level.

$$B_{ihd} = \max(T_{hd} - b_i, 0)$$
$$C_{ihd} = \max(T_{hd} - c_i, 0)$$ (2)

where T_{hd} is the average hourly temperature of day d, and b_i and c_i are the unit-specific breakpoints. These equations are similar in structure to the calculation of cooling degree days, except that cooling degree days is calculated with a set breakpoint, usually 65°F, whereas here we allow the breakpoints to take a range of values. α_i, γ_i, δ_i, η_i, θ_i, ζ_i, b_i and c_i are the parameters to be estimated.

For each unit, we implement a grid search to find the breakpoint values that fit the consumption patterns best, using only non-curtailment weekdays. We estimate the model for values of b_i in the range of 65°F to 84°F and for values of c_i in the range of 78°F to 90°F, both in increments of 1° and with the constraint that $b_i < c_i$. The optimal combination of b_i and c_i is chosen based on which model maximizes R-squared.[5]

One drawback to using the regression-based approach is that out-of-sample predictions could perform poorly due to the parameterization. Regression-based baseline models assume that the functional relationship between load and weather (cooling degree days) is the same on curtailment days as on the non-curtailment days. There are two reasons to be optimistic that this is not a large problem in our case. First, our empirical approach allows for non-linearity in the load-weather relationship, which should enable the upper tail of the distribution to be modeled better. Second, as explained in more detail in the next section, curtailment days were not temperature outliers, and so our out-of-sample weather has strong statistical overlap with the in-sample days.

From the parameters of the baseline models determined with non-curtailment days, we then predict demand on curtailment days. The difference between actual consumption and predicted consumption is the key statistic for determining the effect of the program for both hours during the event and hours before and after the event.

---

[5] While R-squared is typically not thought to be a useful judge of predictive power, in this case we are only comparing models with the same number of explanatory variables. So R-squared would indicate the same breakpoints as optimal as would AIC or BIC.

4 Data

This section details the multiple sources of data used for the analysis. The main dataset is the 15-minute interval consumption data covering the period June 12 through October 31, 2007. We additionally use weather data to model consumption and data on the marginal cost of generating 1 MWh in the wholesale market to monetize the benefits of changes in consumption on curtailment days.

4.1 Electric Load Data

As part of the SmartAC program load impact assessment, AC load loggers were installed on a sample of 352 participating units. These units were contained in a stratified random sample of 297 (146 PCTs and 151 switches) program participants from a population 2,917.[6] The sample was stratified by type of device (PCT or switch), total number of cooling tons from all units (less than four tons, four tons or more), and multiple units (yes or no). The sample design, and stratification as well as design sample sizes, and installed sample sizes are presented in Table A1 in the Appendix. We removed one unit for having an anomalously high load of 253 kW (over 10 times larger than the mean load) and four units for having 0 kW consumption throughout the study period. 53 units were dropped because of missing data, which left us with a final sample of 254 participants and 294 loggers.

Electricity consumption for the participating units were recorded by the data loggers in 15-minute ampere readings, which recorded the average instantaneous amp reading for each minute during that period.[7] The amperage was converted to kW using the voltage levels measured at the units. The average instantaneous readings were then converted to kWh by summing each participating unit's kW consumption over the four 15-minute interval readings within that hour. We discarded data on weekends and holidays, as curtailment days are only called on normal weekdays. In total, we had over 3.9 million observations at the 15-minute-unit level. Mean consumption was 0.26 kW with a standard deviation of 0.85, while the maximum kW value was 7.66.

---

[6] Some participants had more than one unit.

[7] We do not have data on non-AC electricity consumption. If households change consumption on other margins, such as using fans more, we hypothesize that this will make our estimates of consumption changes during curtailment larger (more negative) than the total changes in electricity, but is unlikely to affect non-curtailment times.

4.2 Temperature Data

      Hourly temperatures and relative humidity were obtained from NOAA National Climatic Data Center. While we do not know the location of each sample unit, PG&E mapped each unit to the nearest weather station, and we use this to match each sample unit with weather data.

      Figure 2 is a plot of the density of the average daily temperature data used in the analysis with vertical lines representing the average daily temperature at which curtailment events were called. Due to the experimental nature of this program in 2007, many curtailment days were called for average heat days when peak load reductions were unlikely needed. This is a critical facet of our study because it ensures that our out-of-sample predictions will have strong statistical overlap with the in-sample days.

4.3 Marginal Cost of Generating Electricity

      In order to monetize the net savings from the SmartAC program, we obtained data on the marginal cost of generating 1 MWh of electricity from California ISO (CAISO).[8] We use the locational marginal pricing (LMP) at Stockton, CA. We believe that the conditions existing in 2007 in CAISO means that the price data we use are indicative of both the relative scarcity and abundance of the product and the sensitivity of the demand to variations in price. The CAISO market uses a balancing energy market in allocating energy to match load less than 10 minutes ahead forecasting scenarios. This market has two important features, namely an hourly, paid-as-bid market that is conducted at the intertie points with neighboring control areas, and a five-minute, single price auction market for internal generators. Prices for the balancing energy market are settled at a market determined price that is set by the most expensive energy dispatched by the system in any given interval using a uniform price auction. The LMP reflects the energy value at a specific location at a specific time, fluctuating with time of day, and can surge significantly with extreme temperature and other unpredictable events. The LMP data we use is hourly. During peak load hours on curtailment days, the mean price is $100/MWh, with a maximum of $230/MWh and a minimum of $60/MWh. During off-peak times, the mean LMP value is $58/MWh.

---

[8] http://www.ferc.gov/market-oversight/mkt-electric/california/caiso-archives.asp

5 Results

5.1 Baseline model

While our primary interest is changes in consumption patterns induced by curtailment days, we first examine the baseline model estimation for indications that it captures demand well. It terms of model fit, the average R-squared across all 294 units is 0.76 with a 90% confidence interval of [0.53, 0.97]. This is an improvement over papers using temperature regressions or single breakpoint models (Herter 2007, Reed et al. 2009, Newsham et al. 2011). Further, when also accounting for between-unit variation, our model captures over 94% of the total variation in cooling load.

The breakpoints resulting from the optimization showed substantial heterogeneity.[9] Figures 3a and 3b are density plots of the estimated first and second breakpoints, respectively, for participating AC units. The mean value of the first breakpoint is 71°F with a 90% confidence interval of [66, 82], which seem like reasonable temperatures for AC units to be switched on. The distribution of the second breakpoint appears bimodal, with peaks at 80°F and 88°F. The two figures illustrate how different units have different breakpoints because of the differences in the buildings mass, outside temperature and user preferences. This indicates the importance of modeling this heterogeneity, which may be lost with a more traditional panel estimator.

In addition, the coefficient estimates on the breakpoint parameters were highly statistically significant. 88.4% were significant at the 0.1% level, and the remaining were significant at either 5% or 10% level. In contrast, when the models are estimated with uniform base temperature values of 65°F and 72°F for the first and second breakpoints, respectively, slightly less than 65% of the coefficient estimates were significant at the 10% level. Thus, our grid search method substantially improves model fit.

5.2 Changes in consumption on curtailment days

Figure 4 illustrates our key results for a single curtailment day. The solid line represents the predicted load profile averaged across all units, the dashed line is the actual load, and the vertical lines mark the beginning and end of the curtailment period. The figure shows that actual consumption rises above predicted before curtailment begins, falls during the event, and rises

---

[9] For less than 3% of the units, the optimal model fit did not include a second breakpoint. All units had the first breakpoint in the baseline model.

again after the event. There are clear savings during peak load, but some of these savings are offset by shifts in the load.

Table 1 presents the full results of our analysis of estimated change in consumption before, during and after a curtailment event on a curtailment day. Results are presented for each of the 13 curtailment days as well as an overall mean. We normalize kWh reductions per hour because the length of curtailment varies between four and six hours. The values in parentheses are the estimated standard deviations of mean change in consumption. Using this measure, all but a few estimated changes are statistically different from zero.

During the curtailment event, the results indicate that the average sample AC unit reduced load between 0.020 kWh/h to 0.754 kWh/h, with an average reduction of 0.410 kWh/h across all curtailment days. This equates to a 20.9% average reduction in peak load consumption for participating units. Peak load reductions tended to be larger on hotter days. These estimates are consistent with prior studies that have estimated the load impact of DR programs and found values ranging from 0 kWh/h to 0.75 kWh/h (KEMA 2008, Herter et al. 2007, Anthony 2009).

Table 1 also shows increases in consumption before and after the curtailment period. Our results suggest that precooling occurred in 11 of the 13 curtailment days, ranging from 0.040 kWh/h to 0.320 kWh/h with an average value of 0.127 kWh/h.[10] Snapback ranges from 0.009 kWh/h to 0.240 kWh/h with mean of 0.115 kWh/h. The magnitude of both precooling and snapback increases with hotter days. Relative to normal AC load, our estimates suggest a 13.9% increase in load prior to an event and a 13.7% increase in load after an event. On a per hour basis, pre-cooling equates to 31% of the reduction during the curtailment period, and snapback equates to 28%. These results quantify behavior taken by households to maintain comfort during curtailment days.

5.2.1 Heterogeneity by technology and load

We are also interested in how curtailment affects units with different installed technologies (PCT versus switch) and units with different loads. Table 2 shows the average changes in consumption averaged across all curtailment days splitting the sample various ways.

---

[10] The before and after periods were defined as three-hour windows on either side of the curtailment period. A three hour period was chosen because it tends to capture all relevant changes in consumption. This was empirically determined by estimating changes in consumption for every hour of every curtailment day, shown in Table A1 in the online appendix.

Panel A splits the sample by choice of technology. The results show the same general pattern of decreases in consumption during the event and increases in consumption before and after the event for both groups. However, the changes for the PCT group are larger in every stage – more reductions during the event and more pre-cooling and snapback. One possible explanation for the heterogeneous change in consumption between the PCT units relative to those with adaptive switches could be related to how the two technologies effect load reduction during curtailment days. Another reason could be that the larger changes in consumption by thermostat compared to the adaptive switches is a reflection of the higher variation in thermostat load compared to switches. The algorithms for the two technologies are such that load reduction by units with adaptive switches is predictable and more evenly distributed over the curtailment period compared to units with PCTs.

Also, within the PCT group, the two different strategies for achieving load reduction result in different outcomes during the curtailment events, but do not seem to affect the amount of pre-cooling and snapback. PCT units, which are exposed to the steep temperature declines, reduce about 13% more load on average.

Panel B partitions units into quartiles based on load size.[11] Similar to Panel A, we see the same pattern across all four quartiles. However, another pattern emerges that as load grows so does the magnitude of consumption change. This finding is consistent with intuition because units with a larger load are likely to be used more and thus have larger decreases during curtailment. Further, larger units are likely to pre-cool more if occupants want to maintain the same comfort level during the curtailment event.

5.2.2 Falsification test

In order to alleviate concerns that the estimated changes in consumption during event days were spurious results, we conducted a falsification test with proxy days treated as curtailment days. To implement this, we sought a match for each curtailment day that had the same average temperature, was within two days of the curtailment day, was not a curtailment day itself, and was not a weekend or holiday. Matches meeting these criteria were found for eight of 13 curtailment days. We then re-estimated our baseline model excluding both the true

---

[11] Figure A2 in the online appendix plots the density of load size across units and the quartile positions used for this split. About fifty percent of the participating units' rated kW loads were less than 20 kW with the lowest unit having a rated load of 2.8 kW and the maximum unit with rated kW load of 35 kW.

curtailment days and the proxy days, and then calculated the difference between actual and expected consumption for the proxy days. The results, shown in Table 3, suggest no evidence of systematic increases or decreases in consumption during any of the three periods (before, during and after the event). Magnitudes in all three columns tend to be smaller than the numbers in Table 1 and flip sign from day to day.

5.3 Calculating peak load savings and net savings

We now examine the complete effect of the SmartAC program in terms of energy savings, as well as monetary savings and environmental benefits, accounting for the mitigating effect that pre-cooling and snapback have. Table 4 reports the results. Columns A, B and C present the peak load benefits in terms of kWh, $ and CO_2 emissions, respectively, on per unit basis. Columns D, E and F present net benefits for the same measures. kWh savings come directly from our model estimates. Monetary benefits are calculated using changes in consumption multiplied by the hourly cost of generating 1MWh. Graff-Zivin et al. (2014) estimate hourly marginal emissions intensities for electricity generation in the Western Interconnect, and we use these estimates to derive CO_2 impacts.

As expected and as foreshadowed in earlier results, on average, units reduced consumption by a total of 23.2 kWh during the summer's peak events. Monetizing these reductions indicates that PG&E saved each participant $41. Further, these changes in consumption reduced a total of 10.1 pounds of CO_2 equivalent, providing additional external benefit to the program. However, after accounting for precooling and snapback, kWh savings are reduced by 41%, monetary savings are reduced by 42%, and CO_2 benefits are reduced by 46%.[12] These results make clear the importance of accounting for behavioral responses to DLC programs when determining program benefits.

Column G presents the cost of conservation, which is calculated as the marginal cost of generating the extra electricity needed to meet the increased off-peak demand due to conservation during peak times. The values indicate that reductions during peak times cost the utilities an additional $21.07/MWh on average, with costs reaching as high as $61.11/MWh for a particularly hot day. Of course, this is cheaper than the cost of generation during peak times, so

---

[12] The discrepancy between percentage reductions is due to the higher generation costs and different fuels being used during peak times.

the net benefit is still positive. But it is important to know that the cost is not zero. These estimates are quite similar to those of Holladay et al. (2014), who estimate conservations costs ranging from \$43.70/MWh to \$61.70/MWh.

Despite the large reductions in net benefits from pre-cooling and snapback, the program still offers substantial benefits. Projecting the net savings onto the entire 8,193 customers enrolled in the program leads to an estimated \$194,000 in utility bill savings. Further, the hotter days in the latter part of the summer yield the largest monetary savings, due both to larger kWh reductions and higher marginal prices. This result suggests that per-unit savings are likely to be larger for summers in which curtailment occurs mostly on very hot days.

6 Policy Discussion

The purpose of this section is to review policy options to reduce peak load and place our results in the context of these options. We review price incentive programs, behavioral programs and technology only programs.

Programs with price incentives achieve peak load reductions by encouraging either the elimination of on-peak energy consuming activities or shifting of such activities to other periods through economic incentives. The economic incentives are usually in the form of differential pricing in residential electricity rate structure based on time of delivery of electricity. Critical peak pricing (CPP), which increases prices during peak hours only for a few days per year, has been shown to reduce peak loads by 13-20% (Newsham and Bowker 2010). Another incentive scheme is time-of-use (TOU) pricing, which charges differential prices by time of day, but does so for entire seasons. This approach produces smaller changes in consumption, in the range of 3-6%, because it is harder for households to permanently change daily behavior compared to a one-off shift where they can just avoid behaviors for a single day.[13] Peak time rebate (PTR) programs, which pay consumers to reduce consumption during peak load, have also been shown to be less effective than CPP (Newsham and Bowker 2010). This result is in line with evidence from Psychology which found that people respond more to "sticks" than "carrots" (Kahneman and Tversky 1984).

Behavioral based programs do not use monetary incentives, but instead harness psychological factors and provide information to nudge people to consume less. Allcott (2011)

---

[13] Additionally, CPP programs typically have higher peak prices than TOU programs.

finds that providing households with monthly feedback on their consumption relative to neighbors can reduce average energy consumption 2%, with larger effects for households with higher baseline consumption. Jessoe and Rapson (2014) implement a field experiment that examines residential electricity price responsiveness and combines CPP with high-frequency information. They find that households reduced peak demand by 0-7% in response to price incentives alone. However, households that were provided with information on real-time usage and prices reduced peak load by 8-22%.

The use of technology is another strategy that can reduce peak load and energy consumption more generally. Technological solutions can be thought of in two dimensions, energy efficiency (EE) and DLC. EE improvements aim to provide the same energy services using less energy and can include upgrades to appliances or lighting, weatherization and fuel economy standards. Gillingham et al. (2006) estimate that EE investments led to a 12% reduction in energy use in buildings, though there is significant debate about the magnitude of gains available going forward.[14] DLC programs, like the SmartAC program we examine in this article, externally control appliances during peak demand times on a small number of event days. DLC programs have been shown to reduce peak load consumption 22-47% during curtailment periods (Newsham and Bowker 2010, Faruiqui and Wood 2008). Our results of the peak load impacts of a DLC program are on the low end of this spectrum.

With this suite of program options available, how should utilities proceed? The evidence suggests that technology based interventions provide the largest reductions in peak load (DLC) and overall (EE) consumption. Technology based programs exclude many margins of adjustment that price and behavior programs allow, and this is why reductions are larger. Essentially, households have only one decision to make, participate in the program or not, and then reductions are automatic versus programs where decisions need to be made many times a season or even many times a day for reductions to be realized.

However, while peak reductions are large for DLC, our results show that much of these gains can be offset by increases in consumption at other times. In this light, the specific design of

---

[14] Granade et al. (2009), in a report for McKinsey & Company, describe energy efficiency as a vast, low-cost resource worth over $1.2 trillion if upfront investment in efficiency measures through 2020 were executed at scale. However, Allcott and Greenstone (2012) argue that the gap is likely very small, in the order of 1-2% of energy use. The discrepancy is often attributed to differences in engineering versus economic estimates of gains (e.g., Dubin et al. 1986, Nadel and Keating 1991), though recent evidence suggests engineering estimates may be improving and gains are real (Lang and Siler 2013).

Jessoe and Rapson (2014) that combined CPP with high frequency information may be optimal because peak reductions were almost as large as with SmartAC and there were no increases in consumption at other times of day. In fact, in stark contrast to the negative temporal spillovers we that we find, Jessoe and Rapson find positive spillovers (reduced consumption) on non-event days, suggesting that through CPP and information households are learning to manage demand.

7 Conclusions

Deregulation has greatly impacted the ability for the electricity industry in the US to make profit. Specifically, additional generation capacity is no longer guaranteed a rate of return. Because of the costliness of new capacity and the growing pressure peak demand places on the system, there has been increasing implementation of DR programs as a way to curb peak load. DR can also be applied for energy conservation in general. Recent literature has quantified the benefits of DR by measuring the amount of load reduced during peak periods, and the impacts are substantial. However, little work has been done to quantify the impact of behavioral responses to such programs. This paper seeks to contribute to that gap in the literature by examining the strategic response of participants in one such DLC demand response program and how that response may undermine its effectiveness.

Our results suggest that program participants increase consumption before and after event hours relative to the counterfactual. These results are intuitive, but are also of critical importance as DR policies move forward and the need to understand net benefits and forecast demand grows. For example, the SmartAC program, now in its seventh year, has over 100,000 enrolled customers. To be clear, the results still support DR and DLC programs as an effective way to curb peak demand, but the true conservation cost is not zero.

REFERENCES

Allcott, H., 2011. Social norms and energy conservation. Journal of Public Economics 95, 1082–1095.

Allcott, H., Greenstone, M., 2012. Is there an energy efficiency gap? National Bureau of Economic Research.

Anthony, A.W., 2009. Climate change and peak demand for electricity: Evaluating policies for reducing peak demand under different climate change scenarios.

Berkhout, P.H.G., Muskens, J.C., Velthuijsen, J.W., 2000. Defining the rebound effect. Energy Policy 28, 425–432.

Binswanger, M., 2001. Technological progress and sustainable development: what about the rebound effect? Ecological economics 36, 119–132.

California ISO, 2014. 2013 Annual Report on Market Issues & Performance.

Carrie Armel, K., Gupta, A., Shrimali, G., Albert, A., 2013. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy 52, 213–234. doi:10.1016/j.enpol.2012.08.062

Crew, M.A., Fernando, C.S., Kleindorfer, P.R., 1995. The theory of peak-load pricing: a survey. Journal of regulatory economics 8, 215–248.

Davis, L.W., Kahn, M.E., 2010. International trade in used vehicles: the environmental consequences of NAFTA. American Economic Journal: Economic Policy 58–82.

Dubin, J.A., Miedema, A.K., Chandran, R.V., 1986. Price effects of energy-efficient technologies: a study of residential demand for heating and cooling. The Rand Journal of Economics 310–325.

Energy Information Administration, 2014. Annual Energy Outlook with Projections 2014. Government Printing Office.

Faruqui, A., George, S.S., 2002. The value of dynamic pricing in mass markets. The Electricity Journal 15, 45–55.

Faruqui, A., Wood, L., 2008. Quantifying the Benefits Of Dynamic Pricing In the Mass Market. Prepared by The Brattle Group for Edison Electric Institute.

Fowlie, M.L., 2009. Incomplete environmental regulation, imperfect competition, and emissions leakage. American Economic Journal: Economic Policy 1, 72–112.

Gillingham, K., Newell, R., Palmer, K., 2006. Energy Efficiency Policies: A Retrospective Examination. Annual Review of Environment and Resources, 31:193-237

Gillingham, K., Kotchen, M.J., Rapson, D.S., Wagner, G., 2013. The rebound effect is overplayed. Nature 493, 475–476.

Goulder, L.H., Stavins, R.N., 2011. Interactions between state and federal climate change policies. In The Design and Implementation of US Climate Policy, NBER Chapters, pp. 109-121. National Bureau of Economic Research, Inc.

Graff-Zivin, J., Kotchen, M., Mansur, E.T., 2014. Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and Other Electricity-Shifting Policies. Journal of Economic Behavior and Organization, 107A, 248-268.

Granade, H.C., Creyts, J., Derkach, A., Farese, P., Nyquist, S., Ostrowski, K., 2009. Unlocking Energy Efficiency in the U.S. Economy. McKinsey & Company.

Guerra Santin, O., Itard, L., Visscher, H., 2009. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy and Buildings 41, 1223–1232. doi:10.1016/j.enbuild.2009.07.002

Guerra-Santin, O., Itard, L., 2010. Occupants' behaviour: determinants and effects on residential heating consumption. Building Research & Information 38, 318–338. doi:10.1080/09613211003661074

Herter, K., 2007. Residential implementation of critical-peak pricing of electricity. Energy Policy 35, 2121–2130. doi:10.1016/j.enpol.2006.06.019

Herter, K., McAuliffe, P., Rosenfeld, A., 2007. An exploratory analysis of California residential customer response to critical peak pricing of electricity. Energy 32, 25–34.

Holladay, J.S., Price, M., Wanamaker, M., 2014. The Perverse Impact of Calling for Energy Conservation. Working paper.

Jessoe, K., Rapson, D., 2014. Knowledge is (less) power: Experimental evidence from residential energy use. American Economic Review 104(4), 1417–1438.

Kahneman, D., Tversky, A., 1984. Choices, values, and frames. American psychologist 39, 341.

KEMA Inc., 2008. Final Report Pacific Gas and Electric SmartAC Load Impact Evaluation.

Lang, C., Siler, M., 2013. Engineering estimates versus impact evaluation of energy efficiency projects: Regression discontinuity evidence from a case study. Energy Policy 61, 360–370.

Mathieu, J.L., Callaway, D.S., Kiliccote, S., 2011. Variability in automated responses of commercial buildings and industrial facilities to dynamic electricity prices. Energy and Buildings 43, 3322–3330.

Matsukawa, I., 2001. Household response to optional peak-load pricing of electricity. Journal of Regulatory Economics 20, 249–267.

Nadel, S. M., Keating, K. M. 1991. Engineering estimates vs. impact evaluation results: How do they compare and why? Washington, DC: American Council for an Energy-Efficient Economy.

Newsham, G.R., Birt, B.J., Rowlands, I.H., 2011. A comparison of four methods to evaluate the effect of a utility residential air-conditioner load control program on peak electricity use. Energy Policy 39, 6376–6389.

Newsham, G.R., Bowker, B.G., 2010. The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. Energy Policy 38, 3289–3296.

North American Electric Reliability Corporation (NERC), 2013. 2013 Summer Reliability Assessment.

Reed, D., George, S., Schellenberg, J., Malaspina, P., Berghman, D., 2009. Ex Post Load Impact Estimates for SCE's Demand Response Programs: Agricultural and Pumping Interruptible Program Real Time Pricing. Program Evaluation Report Prepared for Southern California Edison.

Thomas, B.A., Azevedo, I.L., 2013. Estimating direct and indirect rebound effects for U.S. households with input-output analysis Part 1: Theoretical framework. Ecological Economics 86, 199–210.

Wolak, F.A., 2011. Do residential customers respond to hourly prices? Evidence from a dynamic pricing experiment. The American Economic Review 83–87.

Figure 1: Description of the Two Breakpoint Model
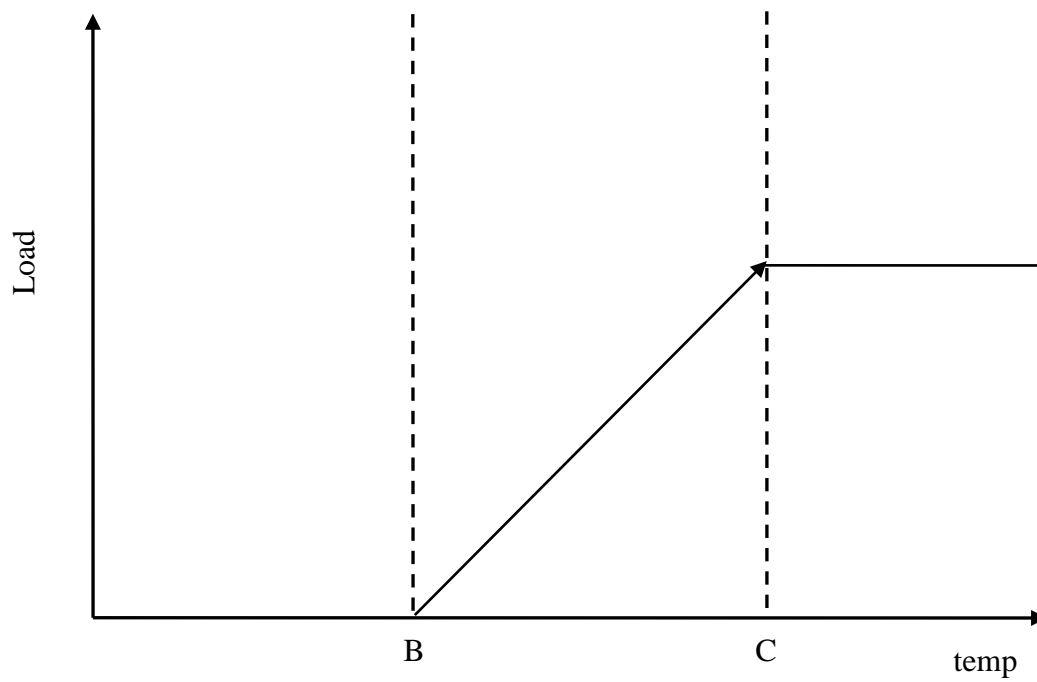
Notes: Point B on the horizontal axis represents the first breakpoint, the point at which cooling starts. Point C, represents the point at which air conditioner runs at 100% duty circle.

Figure 2: Density of Average Daily Temperature Values and Curtailment Days
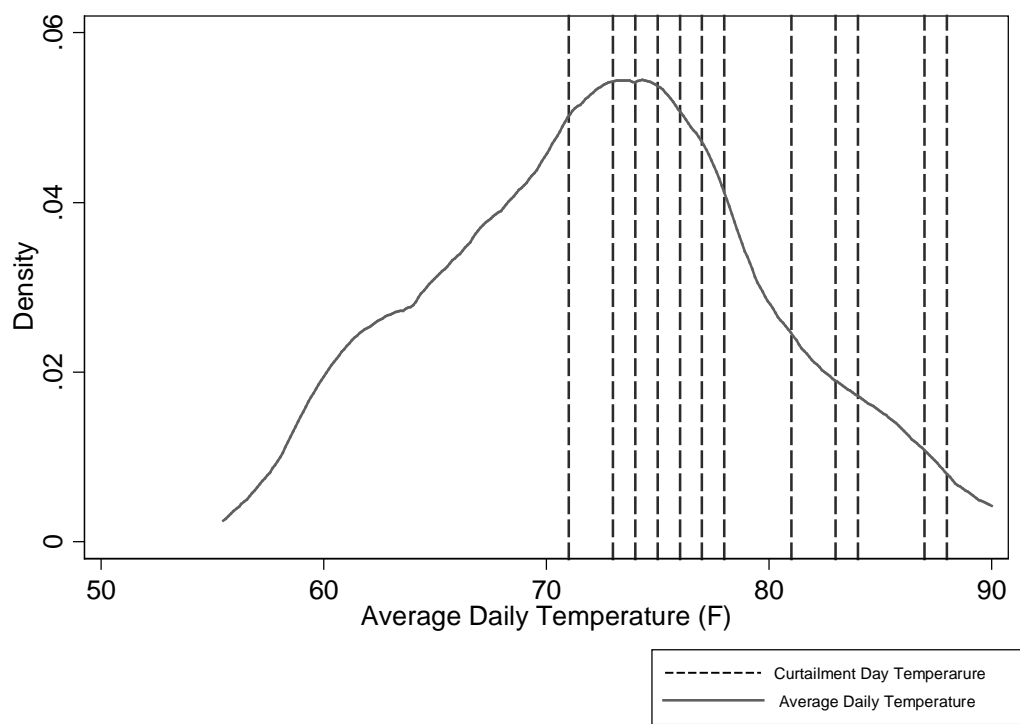
Figure 3a: Distribution of Estimated First Breakpoint Parameter
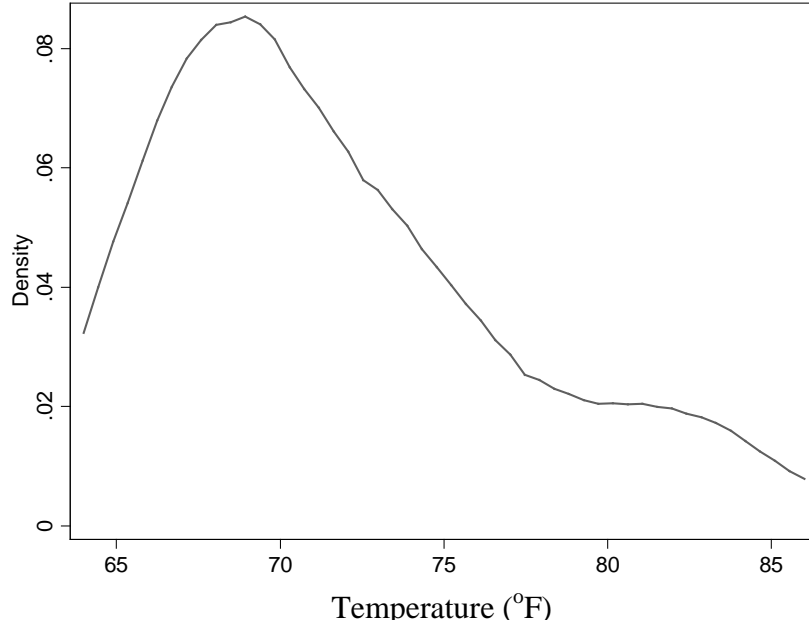
Figure 3b: Distribution of Estimated Second Breakpoint Parameter
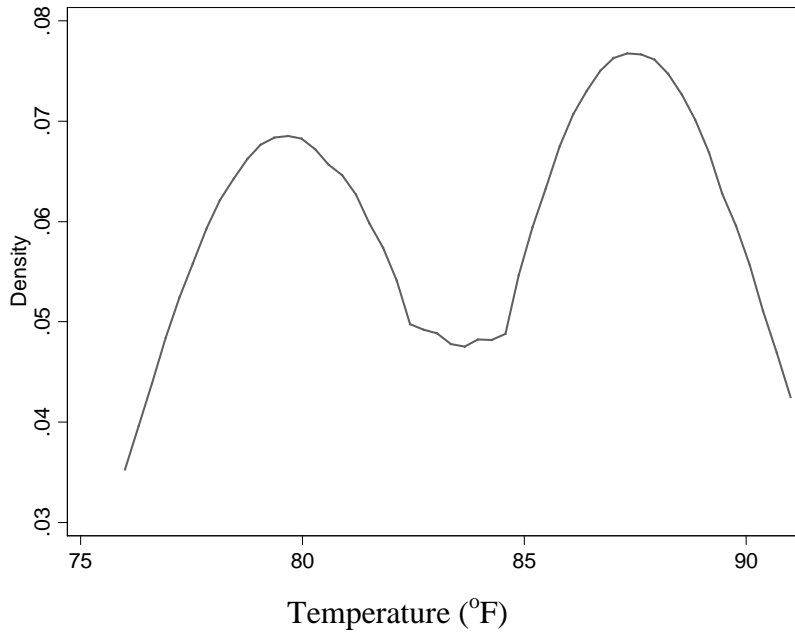
Notes: Figure 3a is a kernel density estimate of the first breakpoint for all the participating AC units, while Figure 3b is the kernel density estimates for the second breakpoint temperature for all participating units. Both densities were estimated with the epanechnikov kernel and optimal smoothing constant of 2.0

Figure 4: Curtailment Day Observed Vs. Predicted Load-Aggregate for August 31
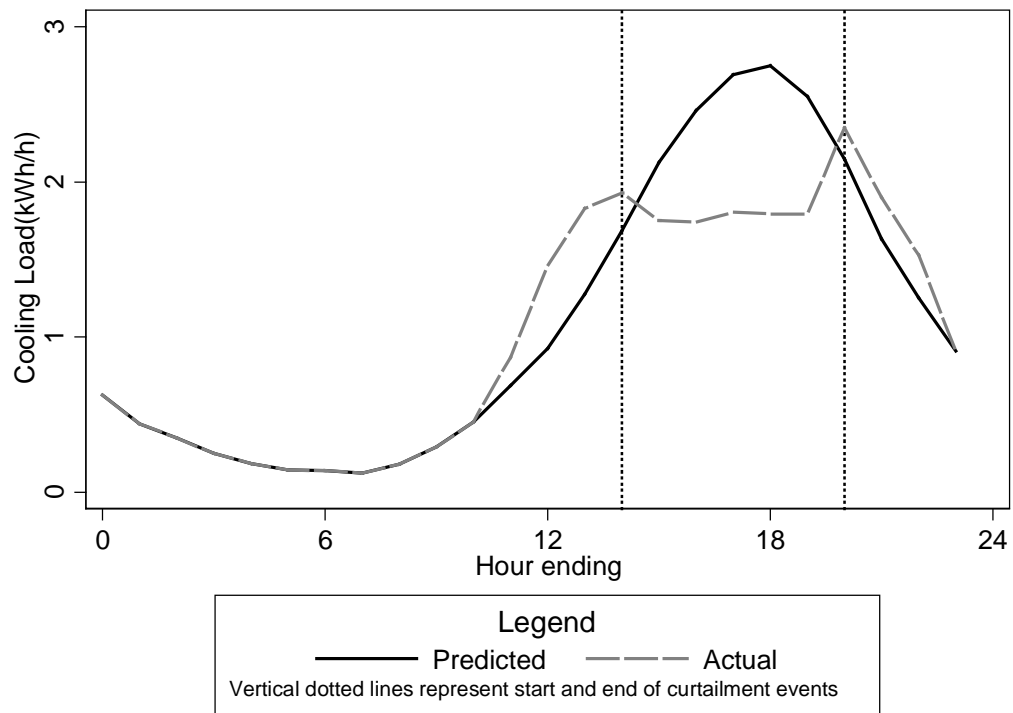
Notes: The predicted load was estimated for each unit, first by estimating the unit specific preferred regression model without the curtailment days. Then we predicted load for each 15-minute interval separately for each unit before averaging across all units for each hour.

Table1: Average Precooling, Load Impact Response & Rebound During Curtailment Days by all Units

Date	Average Daily Temperature (^{o}F)	Average Change in Consumption (kWh/h)		
		Before Event	During Event	After Event
12-Jul	71	0.000	-0.331	0.045
		(0.000)	(0.004)	(0.011)
17-Jul	73	-0.090	-0.020	0.050
		(0.300)	(0.006)	(0.013)
23-Jul	81	0.320	-0.720	0.090
		(0.002)	(0.013)	(0.021)
26-Jul	74	0.040	-0.120	0.109
		(0.001)	(0.001)	(0.015)
27-Jul	76	0.029	-0.140	0.119
		(0.005)	(0.018)	(0.017)
1-Aug	77	0.110	-0.615	0.177
		(0.052)	(0.012)	(0.020)
9-Aug	75	0.060	-0.217	0.009
		(0.003)	(0.013)	(0.002)
10-Aug	78	0.099	-0.320	0.240
		(0.004)	(0.001)	(0.016)
21-Aug	84	0.221	-0.564	0.101
		(0.001)	(0.030)	(0.020)
22-Aug	83	0.159	-0.487	0.180
		(0.001)	(0.001)	(0.021)
28-Aug	84	0.173	-0.553	0.120
		(0.038)	(0.001)	(0.022)
30-Aug	87	0.310	-0.515	0.129
		(0.000)	(0.001)	(0.023)
31-Aug	88	0.230	-0.754	0.133
		(0.000)	(0.034)	(0.022)
Mean		0.127	-0.410	0.115

Notes: The values in columns 3, 4, and 5 were computed by averaging the difference between the predicted mean hourly demand and observed demand on curtailment days. The baseline model was estimated using hourly observations on non-curtailment days alone. Impacts are calculated for each 15-minute interval. The curtailment day load estimates is the same for each 15-minute interval in the hour. The impacts are calculated separately for each observed 15-minute average. The average precooling and rebound values were computed by averaging the differences between the observed and predicted cooling loads in the three hours prior to and after the curtailment event respectively. Standard errors are in parentheses. The standard errors were computed by dividing the standard deviation obtained from a t-test of the difference in means between predicted and observed loads with the number of observations.

Table 2 : Heterogeneity in Response to SmartAC Program

	Average Peak Load on non-curtailment days	Average Change in Consumption (kWh/h)		
		Before Event	During Event	After Event
Panel A: Technology				
PCT Group	2.41	0.190	-0.370	0.273
		(0.001)	(0.001)	(0.001)
Switch Group	1.83	0.090	-0.310	0.140
		(0.000)	(0.001)	(0.001)
Panel B: Classification by Load Size				
1st Quartile	0.68	0.069	-0.150	0.090
		(0.032)	(0.005)	(0.029)
2nd Quartile	1.83	0.114	-0.261	0.073
		(0.001)	(0.005)	(0.002)
3rd Quartile	2.56	0.150	-0.470	0.119
		(0.003)	(0.006)	(0.005)
4th Quartile	2.67	0.318	-0.570	0.191
		(0.020)	(0.007)	(0.001)

Notes: Panel A reports the load impact response by technology type and panel B reports the load impact response by load size in kWh. The reported values are mean values averaged over time and specific class. Standard errors are in parenthesis .

Table 3: Average Precooling, Load Response and Rebound on Proxy days

Date	Change in Consumption (kWh/h)		
	Before Event	During Event	After Event
11-Jul	0.020	-0.010	0.000
	(0.050)	(0.031)	(0.001)
18-Jul	0.009	-0.008	0.000
	(0.017)	(0.019)	(0.009)
19-Jul	0.006	-0.007	0.009
	(0.021)	(0.039)	(0.06)
25-Jul	-0.000	-0.010	-0.047
	(0.006)	(0.074)	(0.06)
31-Jul	-0.000	0.000	-0.010
	(0.001)	(0.00)	(0.03)
2-Aug	0.065	0.000	-0.010
	0.036)	0.000)	(0.053)
27-Aug	0.000	-0.010	0.000
	(0.000)	0.006)	(0.000)
29-Aug	0.010	0.010	0.000
	(0.008)	(0.005)	(0.000)

Notes: Proxy event days are defined as days with similar temperature profile as the curtailment days; and also fall within a few days of the curtailment days (1 or 2 days before or after) the curtailment days that is not a weekend, public holiday, past curtailment day. Model results are calculated for all the curtailment days, however, diagnostics are calculated only for the set of proxy event days.

Table 4: Net Savings with and without Precooling and Snapback

Date	Per Unit Peak Savings			Per Unit Net Savings			Conservation Cost ($/MWh)
	kWh (A)	$ (B)	lbs of CO_2e (C)	kWh (D)	$ (E)	lbs of CO_2e (F)	(G)
12-Jul	1.156	0.724	0.031	1.036	0.339	-0.053	7.24
17-Jul	0.140	0.730	0.133	0.290	0.154	0.109	-2.65
23-Jul	2.884	2.866	1.379	1.654	1.343	0.832	17.14
26-Jul	0.600	3.219	0.237	0.150	0.812	0.033	23.32
27-Jul	0.564	2.919	0.281	0.114	0.547	0.072	35.14
1-Aug	2.476	3.053	1.065	1.636	1.908	0.667	30.65
9-Aug	0.864	1.604	0.388	0.635	1.564	0.279	11.85
10-Aug	1.283	2.731	0.589	0.263	1.143	0.106	49.62
21-Aug	2.795	4.692	1.283	1.835	3.771	0.816	53.83
22-Aug	2.398	3.855	0.855	1.378	1.952	0.360	24.17
28-Aug	2.224	3.855	1.247	1.354	2456	0.698	58.23
30-Aug	2.084	4.747	0.923	0.764	2.676	0.323	61.11
31-Aug	3.750	5.986	1.709	2.670	4.957	1.200	46.00
Sum	23.218	41.040	10.120	13.798	23.622	5.442	

Notes: Columns A and B report the savings over the curtailment period in kWh and Dollar values respectively. Column B is computed by multiplying kWh savings by the number of hours and the total kW signed up to the program as at August 31 2007 (number of units multiplied by the tonnage per unit) as well as Net Marginal Price at that particular hour in $/MWh. Columns D and E report the savings accounting for the precooling and rebound effects in kWh and dollars respectively. Column D was computed by subtracting the amount of precooling and snapback from total reduction during the curtailment period. Column E was computed the same way as Column B, however the net reduction was used. Columns C and F were computed by multiplying hourly change in consumption with hourly emission factors as reported in literature. Column G is the marginal cost of generation for the changes in off-peak consumption.