AUTOMATED DEMAND RESPONSE SYSTEM PILOT

Final Report

Appendices

Rocky Mountain Institute

Table of Contents

Appendix A
Data Collection and Load Impact Analysis Methodology

This section summarizes the sample design, data development and load impact estimation methodology that underlies the load impact results evaluated for the 2004 and 2005 ADRS pilot.

Sample Design

This section characterizes the ADRS pilot participant population and the control population used in the load impact analysis. For the 2005 load impact analysis, one control population was used (identified as A03) while for the 2004 load impact analysis, two control populations were used (A03 plus and additional control group, identified as A07). Characterization of the 2004 control group A07 is included here given that load impact results from 2005 are compared against load impact results from 2004.

All participant and control homes in the ADRS pilot shared three key characteristics: homes were single family, detached units with central air conditioning located in climate zone 3. ADRS participants and control homes were each stratified into two sub-samples according to average daily consumption or usage (ADU). Homes are designated as high consumption if ADU is greater than or equal to 24 kWh per day. Homes with an ADU of less than 24 kWh per day on average were designated as low consumption homes.

ADRS Participants

A total of 175 homes were initially recruited into the ADRS pilot program in 2004, consisting of 75 homes from PG&E, 76 homes from SCE, and 24 homes from SDG&E. The pilot participants were recruited from owner-occupied, single-family homes from climate zone 3 in neighborhoods served by appropriate cable providers and in zip codes identified by the participating utilities. ADRS homes were recruited at random in 2004 regardless of historical consumption, although homes were screened for eligibility with respect to the presence of central air conditioning and within prescribed zip codes. Because ADRS technology is capable of controlling end uses in the home in addition to central air conditioning, homes were screened for availability of other loads (i.e., swimming pool pumps and spas), but not disqualified from participation in their absence.

ADRS homes were stratified into two sub-samples according to average daily consumption or usage (ADU). For ADRS, stratification was based on monthly billing data from June-September 2003 (summer season), divided by the number of days per month to arrive at an average daily usage. Table 1 breaks down the population of ADRS participants by consumption stratum and by utility at the start of the pilot program in 2004.

Table 1
Count and Distribution of ADRS Homes as of July 1, 2004

	PG&E	SCE	SDG&E
High Stratum	51	71	7
Low Stratum	24	5	17
Total	**75**	**76**	**24**

The ADRS homes used for the 2005 pilot load impact analysis consisted of the same households that remained in the pilot program after the summer of 2004. ADRS participants were notified that the pilot would be extended for an additional year and were promised a $125 incentive payment to be paid out in early November 2005 if they stayed with the program through December 31, 2005. The ADRS program was offered to incoming residents of existing ADRS homes, in the event that a current resident rented or sold their home. However, no additional homes were recruited for the 2005 pilot extension.

However, by the start of the second year of the pilot in June 2005, a number of participants opted out of the program, resulting in the following population sizes by utility (Table 2):

Table 2
Count and Distribution of ADRS Homes as of July 1, 2005

	PG&E	SCE	SDG&E	Total
High Stratum	40	53	6	99
Low Stratum	19	4	9	32
Total	**59**	**57**	**15**	**131**

Control sample design

Control homes all with identification numbers beginning with A03 are a subset of the control homes used in California's statewide pricing pilot (SPP), a pricing-only peak load reduction pilot program that ran concurrently with the ADRS program in 2004. The A03 control homes selected into the ADRS pilot resembled ADRS participants in three key parameters: single-family homes in climate zone 3 with central air conditioning. Consumption stratification of A03 homes was assigned as part of the SPP program using the same convention as for ADRS, in which homes are designated as high consumption if their ADU was greater than or equal to 24 kWh per day, and designated as low consumption otherwise. Table 3 counts the number of A03 control homes extracted from the SPP climate zone 3 population sample at the start of the ADRS pilot period in 2004.

Table 3: Count of A03 Control Homes as of July 1, 2004

	PG&E	SCE	SDG&E	Total
High Stratum	12	22	3	37
Low Stratum	3	14	3	12
Total	**15**	**28**	**6**	**49**

In addition to A03 control homes, a subset of SPP participants on the CPP-F dynamic pricing tariff were used for comparison against ADRS homes in the 2004 load impact analysis. These homes had identification numbers beginning with "A07". Because they were on the CPP-F rate but did not possess ADRS technology, comparison of ADRS load impact relative to this group served as a rough proxy for the incremental impact of ADRS technology.

The A07 homes resembled ADRS participants in three key parameters: single-family homes in climate zone 3 with central air conditioning. Consumption stratification of A07 homes was assigned as part of the SPP program using the same convention as for ADRS, in which homes are designated as high consumption if their ADU was greater than or equal to 24 kWh per day, and designated as low consumption otherwise. Table 4 counts the number of A03 control homes extracted from the SPP climate zone 3 sample at the start of the ADRS pilot period in 2004.

Table 4: Count of A07 Control Homes as of July 1, 2004

	PG&E	**SCE**	**SDG&E**	**Total**
High Stratum	20	38	5	53
Low Stratum	9	16	2	27
Total	**29**	**54**	**7**	**90**

)Augmentation of A03 control population

Load impact evaluation in 2005 shifted emphasis from statewide reporting of results to reporting results by utility. In addition, the focus of the evaluation shifted to the high consumption homes, given their higher performance during the 2004 pilot period relative to the rest of the pilot participants. Examination of the control sample (A03) revealed that the number of high consumption control homes was too small to make statistically significant inferences about load impact for each utility separately. RMI thus moved to secure additional high consumption control homes for the 2005 ADRS pilot period. Because one of the objectives of the 2005 ADRS pilot includes a comparison of 2005 load impact results with those of 2004, we ultimately re-evaluated the 2004 load impact results of high consumption ADRS homes against the augmented population of high consumption control homes.

The following paragraphs describe our methodology for determining the number of additional high consumption homes needed to achieve statistically significant load impact results. Ordinarily in a statistical experiment, the sample size would be set based on the underlying desired precision and expected standard deviation. The statistical formula for sample size is:

$$Z^2 / H^2 * (\sigma)^2 = N, \text{ where}$$

Z = z-value of standard normal distribution curve corresponding to desired level of confidence
H = desired level of precision
σ = standard deviation of the sample, and
N = sample size

From the prior pilot data, the average standard deviation in average kW consumption for A03 homes is 2.06 kW in the high consumption stratum and 1.45 kW in the low consumption stratum. If the precision is to be +/- 0.55 kW, and we want to target a 90% confidence interval (corresponding to a z-value of 1.65), then the sample size should be:

$$(1.65)^2 / (0.55)^2 * (2)^2 = 36 \quad \text{High consumption homes}$$

Under these circumstances, a 40-home sample would be effective to account for unforeseen circumstances such as homeowners moving or dropped/missing meter data for a given sample at a given time. Note that this low standard deviation is achieved within population strata. By contrast, an entirely random sample of ADRS residential customers from across all housing types would likely result in a higher standard deviation and would not accurately represent the appropriate target consumer of GoodWatts.

At the end of the pilot period in 2004, there were 36 total control homes in the high consumption stratum from all three utilities combines. A total of 40 high consumption control homes from each utility would bring the total control population to 120. However, having a comparable proportion in the number of control homes to ADRS homes by utility is also desirable to facilitate ease of statewide reporting of 2005 results in comparison to 2004 in fulfillment one of the 2005 pilot objectives. Thus, the final recommendation was to augment the control population mostly for PG&E and SCE, with fewer additions for SDG&E to keep the proportional weightings similar.

A total of 40 high consumption control homes from PG&E and SCE, and 10 control homes from SDG&E would bring the total control population to 90, which is then comparable in magnitude to the ADRS population. This means addition of 29 control homes from PG&E, 18 control homes from SCE, and 7 control homes from SDG&E for a total A03 distribution of 40-40-10 (a proportion of 44%-44%-11%, or roughly the ADRS home distribution).

Data collection

The three utilities provided 15-minute interval load data for ADRS and control homes for the period June 1 through October 31, 2005 (June 1 through September 30 in 2004). Because ADRS homes have an additional interval meter as part of the ADRS technology (GoodWatts) package, this second source of load data was also available, and downloadable in real time via the Internet. The ADRS meters serve as a backup to utility meter data as part of the pilot project design in the event that any of the utility meter data were unavailable at the time of the load impact analysis. In the 2004 ADRS load impact analysis, GoodWatts meter data were used in place of SCE data for September[1]. In 2005, no GoodWatts meter data were used, since all data were successfully collected by utilities for all months over the summer period (June-October).

Customer and ADRS pilot load data received from utilities typically contained several customer-days[2] with blank readings or zero readings. RMI screened the data and removed customer days that contained zero readings or blank readings. Typically the customer days with zero or blank reading constituted a very small percentage (less than 1 percent) of the overall data set. In addition, other adjustments were made. SCE and PG&E data were transmitted as kW loads for each 15-minute period. On the other hand, SDG&E data were transmitted as kWh consumption for each 15-minute period. To facilitate analysis, SDG&E data were converted to kW units by multiplying data values by four.

[1]Verification of interval load data recorded by Invensys meters compared to Utility meters was performed in 2004. Results of the verification showing that data difference between the two meters was less than 1 percent is reported in Rocky Mountain Institute's 2004 Load Impact Evaluation Report.
[2] Customer-day is defined as the data set for one customer for one day.

It should be noted that a complete data set for PG&E's low consumption control homes was not available. Data for these homes include only June 1st through September 13th. As a result, any event day results reported for PG&E's low consumption homes reflect only the five event-days in the June through August period.

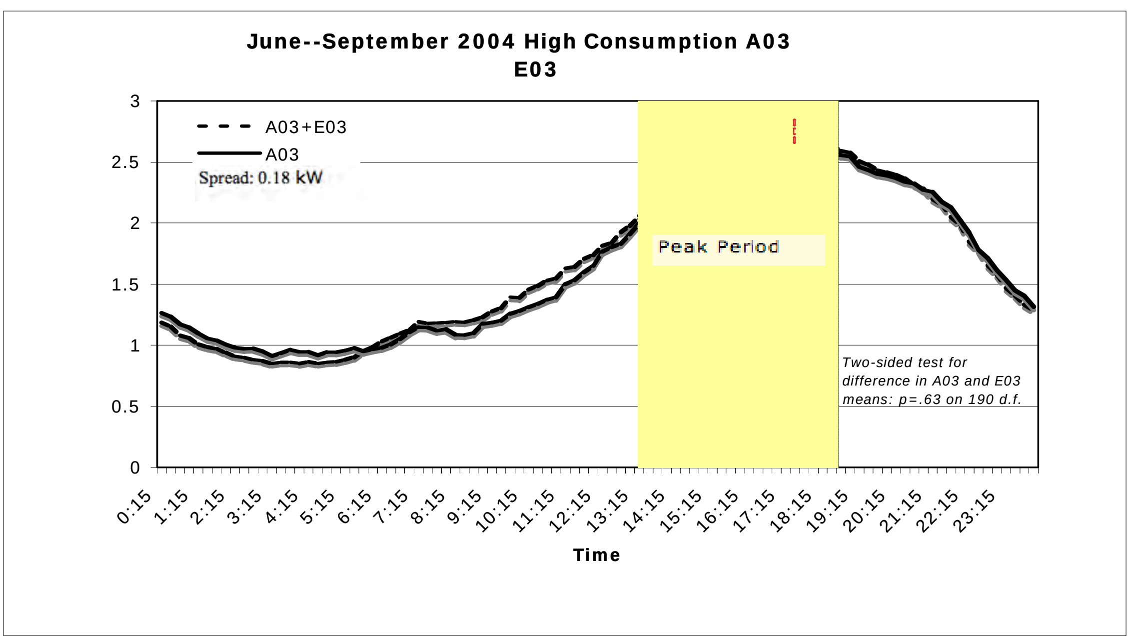
Zip codes and temperature data

Hourly temperature data were collected for June 1- October 31, 2005. Different methods were used for collecting temperature data on control and ADRS homes. For control homes, Invensys provided hourly temperature data through their weather subscription service by weather station, based on home zip code information provided by utilities. For ADRS homes, temperature data were also based on Invensys' weather subscription service by zip code but were downloaded directly from Invensys' GoodWatts website. The temperature data were the same for all ADRS homes for a given zip code. Zip code information for ADRS homes was extracted from the pilot program database administered by Invensys.

Verification of augmentation control homes' load data

In total, the utilities provided 68 augmentation homes: 30 from SCE, 20 from PG&E, and 19 from SDG&E. The additional high consumption control homes thus brought the total high consumption control sample count to 52 for SCE, 32 for PG&E, and 21 for SDG&E. These homes were assigned identification numbers all beginning with the designation "E03" to distinguish them from the original control population. This section presents the results of our verification of E03 load data against the original A03 control sample load data.

Figure 1
Confirmation of E03 Augmentation Control Load Data with A03 Control Data



Supplemental high-consumption E03 homes were compared against the A03 homes in order to determine whether the augmented control group was an appropriate representation of the A03 population. Following detailed investigation, the E03 augmentation sample was accepted and integrated into the A03 sample to form the control sample for use in the 2005 ADRS load impact analysis (see Figure 1). The two-sided test of significance for the difference in average summer loads between the additional and A03 control groups, across all days in 2004, produced a p-value of 0.63. This value indicates a probability of 63% (high) that the differences are due to random chance. For all results presented in this report, "augmented control homes" refers to the control sample that includes the E03 and A03 high consumption homes.

Load impact analysis

To construct average daily loads, RMI averaged the utility interval load data within each 15-minute period across a 24-hour day. Average daily kW loads were calculated for event and non-event days by utility, and by consumption stratum. The averaged daily loads were used to construct event day and non-event day load curves by utility and by consumption stratum. Separate load curves were constructed for ADRS customers, A07 and control customers (A03).

The load curves for the A07 and ADRS customers were then adjusted for selection bias (see discussion in **Error! Reference source not found.**) by adding the appropriate differences adjustments. The difference adjustments, either positive or negative if loads are lower or greater than the control group, respectively, were added to the load curves within each 15-minute data interval. As with the load curves, adjustments were calculated for event and non-event days for each utility, by consumption stratum. For example, the PG&E high consumption ADRS event day load curve was adjusted by adding the PG&E high consumption adjustment. Statewide difference adjustments were calculated from a weighted average of utility-specific difference adjustments. For A07 customers, difference adjustments were made on a statewide basis only, and load impact results are only reported on a statewide basis (see **Error! Reference source not found.**). The quantity of data available for a utility-specific adjustment for A07 consumption was too small and would not have yielded statistically significant results.

ADRS load savings, compared to the control group (A03), were calculated for each 15-minute period by subtracting the adjusted average ADRS load from the corresponding average control home load, for each 15-minute data interval (e.g. PG&E high consumption event day adjusted ADRS loads were subtracted from PG&E high consumption event day control loads). This method is consistent with the "difference of differences" method used by Charles River Associates and California Energy Commission for the larger Statewide Pricing Pilot program. ADRS load reductions were calculated for event and non-event days, by utility and by consumption stratum. The same method was used for calculating ADRS load reductions relative to A07 homes, and for calculating A07 reductions relative to control homes.

Ninety percent confidence intervals were then calculated for average load curves for each 15-minute interval. This range is plotted above and below the mean for a given 15-minute period. Thus we are ninety percent confident that the actual average load of homes in the general population (single family, with central air conditioning, in climate zone 3) are within the range of average load calculated for the sample. By calculating confidence intervals for both ADRS and control homes we also hoped to show that mean differences in load consumption were statistically significant. This was indicated if the confidence intervals above and below the two load curves do not overlap across the peak period.

The ninety percent confidence interval is defined as:

$$\bar{x} \pm 1.645\left(\frac{\sigma}{\sqrt{n}}\right), \text{ where}$$

x bar is the mean for the 15-minute period, σ is the standard deviation of the sample, and n is the sample size. The ±1.645 is the number of standard deviations from a normally distributed mean that contain 90 percent of the sample.

Calculation of peak period reductions

Using the average load reductions calculated for each 15-minute interval on event and non-event days, RMI then calculated Super Peak and peak period reductions for each utility by consumption stratum. Average load drop (kW) across the Super Peak and peak periods was calculated by averaging the load savings curve from 2 p.m. to 7 p.m. on event and non-event days, respectively. The total energy savings (kWh) across Super Peak and peak periods was calculated by summing the 15-minute interval load savings from 2 p.m. to 7 p.m. and then dividing by four[3]. Percentage load reduction during the peak period was calculated by dividing the average load reduction and energy savings during the peak period by the average control load during the peak period.

For the hourly Super Peak and peak period load reductions, the load reduction was averaged for each hour separately from 2 p.m. to 7 p.m. on event and non-event days, respectively. For example, the load savings for each 15-minute period between 2 p.m. and 3 p.m. were averaged to represent the load savings for the 2 o'clock hour. Hourly percent load savings were then calculated by dividing the average hourly load savings by the average hourly control load.

Ninety percent confidence intervals were then calculated for the savings during each hour of the peak period. This was done by first averaging the 90 percent confidence intervals on the control and adjusted ADRS load curves for each hour of the Super Peak period. These errors were then combined to yield a 90 percent confidence interval for the savings (difference between control and ADRS homes) during each hour using standard error propagation techniques.[4]

Exclusion of October data from 2005 load impact analysis

Load impact analysis for the 2005 ADRS pilot was based on performance from July to September only. Although four Super Peak events were called in October 2005, they were excluded from the average load impact calculations for the following four reasons, discussed below.

[3] Because load data were reported in 15-minute intervals, the energy use in any given interval $kWh_1 = kW_1 * (^1/_4 \text{ hr})$. Thus, the energy savings during the peak period is then $(kWh_1 + \ldots + kWh_{20})$ or $(kW_1 + \ldots + kW_{20}) * (^1/_4 \text{ hr})$.

[4] $Error_{combined} = \sqrt{Error^2_{control} + Error^2_{ADRS}}$

First, October events called in the ADRS pilot program are not representative of actual system emergencies during the summer. Typical system emergency events in California occur during the months from July through September, when customer loads reach their annual peak, and when capacity reserve margins are at their lowest as a result.

Second, October events called in the ADRS pilot program are not representative of regional summer temperatures that trigger high demands and actual system emergencies. Figure 2 through Figure 4 show that ADRS homes experienced many days throughout the summers that were hotter than Super Peak days called in 2005. The figures plot the average of maximum temperatures experienced by ADRS homes each day throughout the summer. Black points highlight days when Super Peak events were called statewide.

In PG&E service territory for example, the hottest days were concentrated early in the summer, in July, and declined noticeably by the end of August 2005 (Figure 2). Furthermore, ADRS homes in PG&E's service territory experienced about 32 days in the early summer that were hotter than average peak temperatures on six Super Peak event days called statewide. ADRS homes in SCE service territory experienced very hot days consistently throughout the summer, but there was noticeable decline in temperatures beginning mid-October, after all event days had been called (Figure 3). For ADRS homes in SDG&E territory, it is clear from Figure 4 that temperatures were distinctly different from other regions in the state, with mild temperatures rarely reaching above 90°F. RMI questions the necessity of central air conditioning in homes in climates as mild as those experienced by ADRS homes in SDG&E territory in well-insulated and well-designed homes that are Title24 compliant.

Figure 2
PG&E Summer 2005 Average Daily Peak Temperature

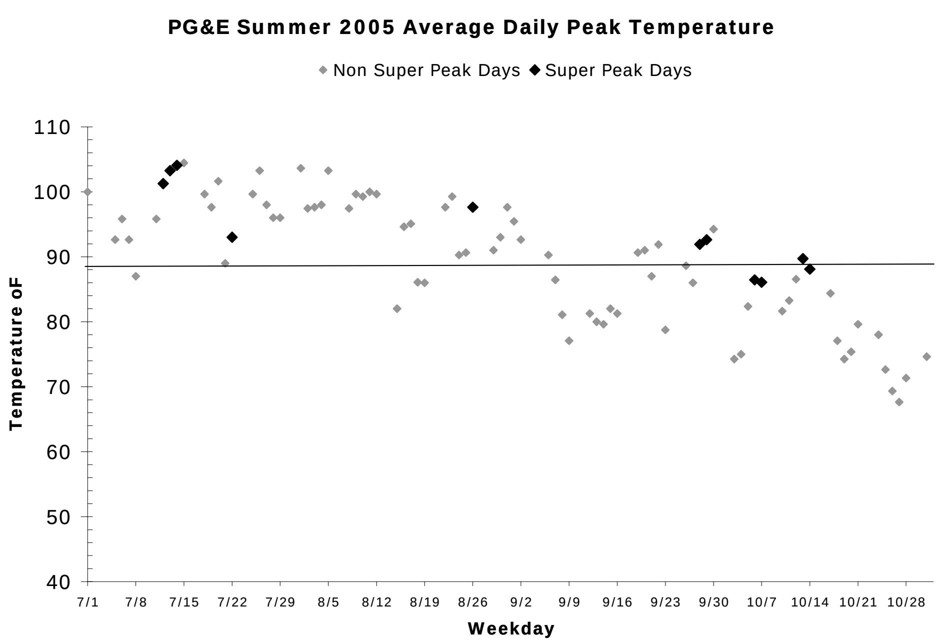


Figure 3
SCE Summer 2005 Average Daily Peak Temperature

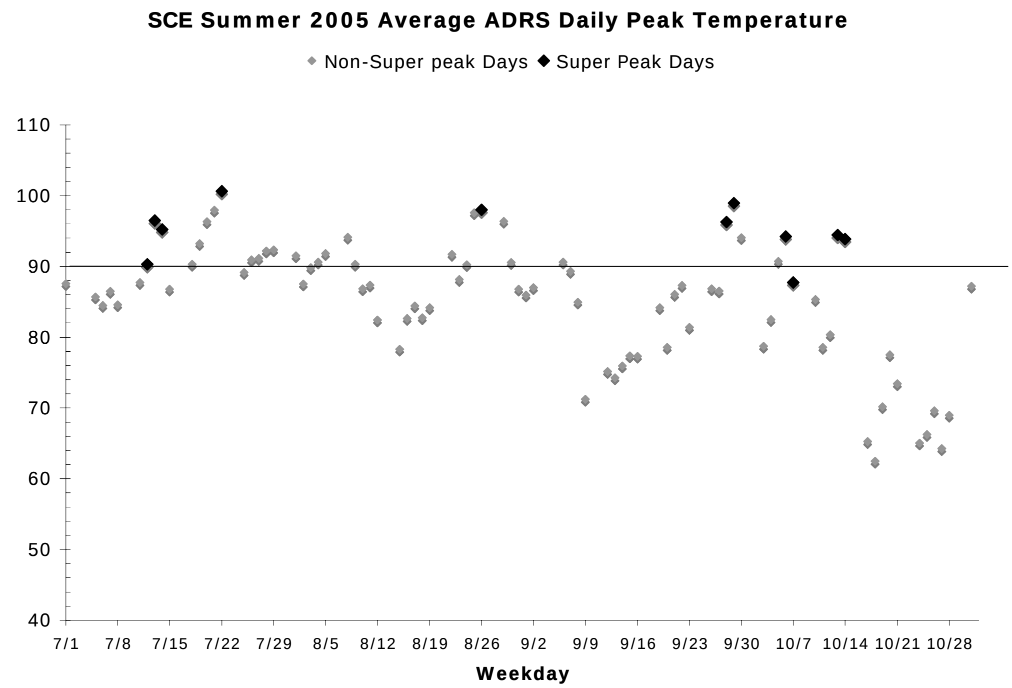
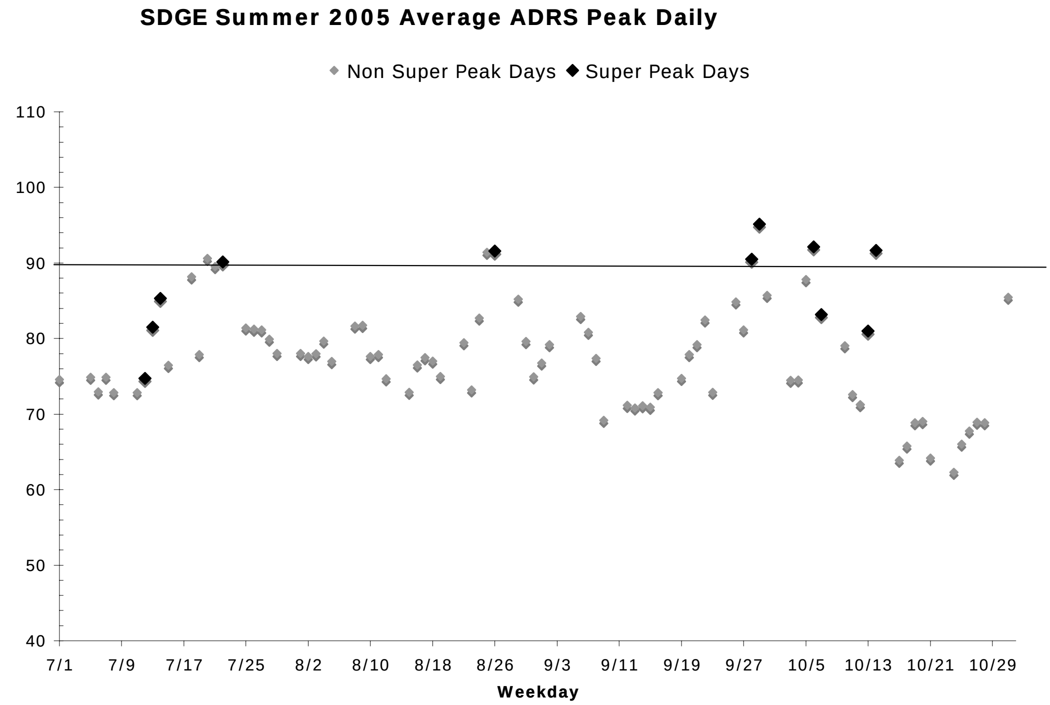


Figure 4
Summer 2005 Average Daily Peak Temperature



Third, October events called in the ADRS pilot program are not representative of insolation values during the interior summer months, July through September. Solar gain is a primary driver of air conditioning load, in addition to temperature. Not only are days noticeably shorter in October, the October sun tends to be much lower in the sky, with associated reductions in solar heat gain inside buildings. Table 5 shows the total solar radiation (beam, diffuse, weather effects) on a horizontal surface such as building rooftops for the summer months in Fresno, CA. Notice that September and October radiation measurements decline to 76 and 60 percent of solar radiation in July, respectively, with associated affects on indoor heat gain and cooling demands.

Table 5
Average Daily Incident Solar Radiation Horizontal Surface (e.g. Roof) for Fresno

	June	July	August	September	October
Solar radiation: Btu/ft^2	2507	2439	2215	1861	1425

Source: Weathermaker v1.01, National Renewable Energy Laboratory, US DOE. 1999.

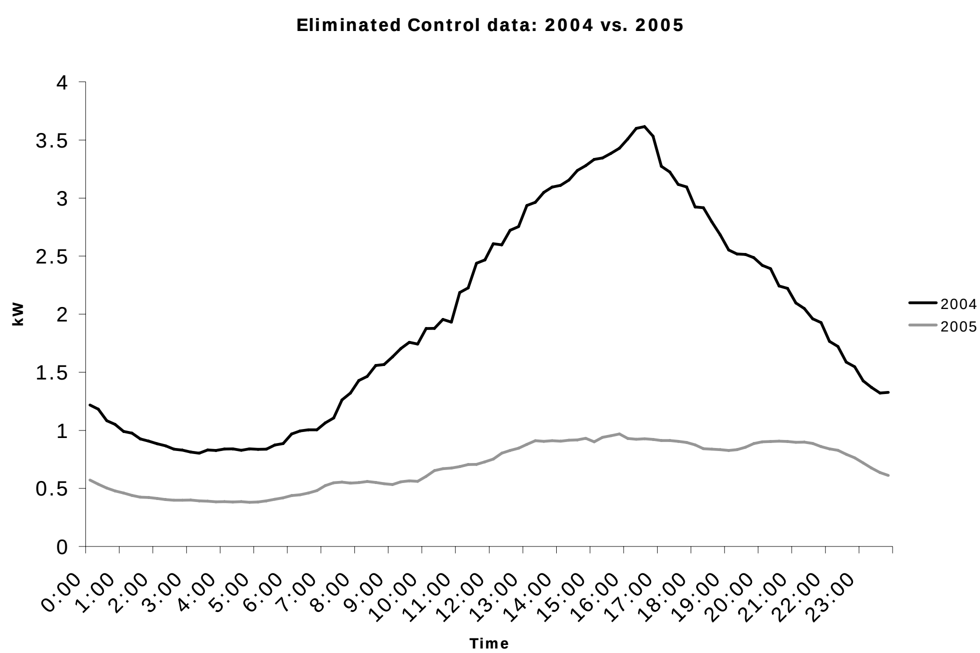
Fourth, October events called in the ADRS pilot program may not be representative of occupancy patterns during the interior summer months. The school year in California starts in September, which potentially drives down total consumption, as children and parents spend more time away from home during the day. Given that October generally tends to be cooler than other times during the summer, ADRS homeowners may change the programming on thermostat settings in anticipation of cooler weather, and use less electricity during the few warmer days that occur in October.

Elimination of outliers in augmented control sample for 2005 analysis

Comparisons of 2004 and 2005 ADRS load reductions revealed that load reductions were lower in 2005 than 2004, but not because ADRS customers consumed significantly more load. Rather, control home loads in 2005 were significantly lower than in 2004. This was in spite of the fact that 2005 was a hotter summer, on average. ADRS loads increased only slightly in 2005, which is consistent with a hotter summer, while control loads decreased, especially during the Super Peak and peak periods.

Delving into the issue in more detail, we examined both 2004 and 2005 control loads for each high consumption control customer in the entire control population. Fifteen-minute interval load data for each high consumption control home were averaged across the entire summer from July through September. Average daily load curves were constructed from the data values and the two years were plotted against each other. We found that generally most control homes' loads were consistent between 2004 and 2005, with some notable exceptions. Nine control homes in SCE and five control homes in SDG&E territories featured normal load profiles in 2004 and nearly flat load profiles in 2005, with consumption in 2005 near zero. Figure 5 shows the 2004 and 2005 average loads for these SCE and SDG&E outlier control houses. Off-peak loads in 2005 were 0.5 kW less than 2004 loads on average and peak period loads differed by as much as 2.7 kW. These 2005 loads seemed to be outliers that were skewing the average control load downward in 2005.

Figure 5: Eliminated control data: 2004 vs. 2005



We adjusted the control homes' data by removing the outlier homes from the summer 2005 control dataset. We did not make the same eliminations for the 2004 data because the control homes in question exhibited normal control home behavior during the 2004 summer. After the elimination of the outlying control home data from the summer 2005 dataset, the 2005 control loads more closely matched 2004 control loads, as seen in the 2004-2005 combined control-ADRS load curves in this report. SCE control loads for 2004 still exceed those for 2005 on event days even after removal of the outlier control homes, but were more similar on non-event days. For SDG&E, 2004 and 2005 control loads were virtually the same on event days after removing the outlying data. On non-event days, 2005 SDG&E loads were greater than 2004 control loads. No outlier control homes were removed from PG&E service territory.

Temperature bin analysis for 2005 pilot

As part of the ADRS load impact evaluation for 2005, RMI was also requested to report ADRS load reductions by temperature bin. This section summarizes the methods we employed for the temperature bin analysis.

First, we grouped temperature data associated with ADRS and control homes into eight temperature bins. Temperatures greater than 85°F were divided into five-degree increments to capture greater detail on hotter days. Temperatures less than 85°F were more coarsely divided into ten-degree increments for evaluating cooler days. These bins are: greater than 105°F, 101-105°F, 96-100°F, 91-95°F, 86-90°F, 76-85°F, 66-75°F, and less than 65°F.

ADRS and control homes within each utility were assigned to one of eight temperature bins. The temperature bin assignment was based on the *maximum* temperature recorded for a zip code associated with an ADRS or control customer on a particular day. We chose this convention on the assumption that the peak temperature experienced drives consumption behavior for ADRS and control customers.

Average kW load reduction was then calculated by temperature bin. Temperature bin results were reported separately for Super Peak and peak periods on event and non-event days, respectively. Consistent with all load impact analyses in this pilot program, all ADRS loads included a selection bias adjustment on event and non-event days, by utility and by consumption stratum (see selection bias discussion in **Error! Reference source not found.**). Instead of applying the adjustment for each 15-minute interval, however, we calculated an average adjustment for the Super Peak and peak period from 2 p.m. to 7 p.m. This selection bias adjustment was then applied to the peak period ADRS load reductions relative to the control sample.

A percentage savings was also calculated for each temperature bin for each consumption stratum, for both event and non-event days. We divided the respective kW savings for each temperature bin by the average control home load in each temperature bin to determine the percentage savings.

Household level analysis

To assess the relative performance of ADRS customers at the household level, RMI calculated the average initial load drop between 2:00 p.m. and 2:15 p.m. and between 2:15 p.m. and 2:30 p.m. for each ADRS customer during every weekday from June 2004 through September 2005. The two time intervals were chosen because the first half hour of the 2:00 p.m. to 7 p.m. peak period generally produces the largest load drop every day. The load drop was calculated as a percent reduction from the period immediately prior, from 1:45p.m. to 2:00 p.m. The larger of the two values, percent load drop during the first fifteen minutes and percent load drop during the second fifteen minutes of the peak period, was used as the representative performance value for each ADRS participant.

The performance values calculated for each day for each ADRS customer were then assigned one of three numeric scores. A "3" score (high performance) was assigned if the average initial load drop was greater than thirty percent. A "2" score (medium performance) was assigned if the average load drop was between twenty and thirty percent. Finally, a "1" score (low performance) was assigned if the initial load drop was less than twenty percent. The daily scores were then segregated by event and non-event days. Finally, average event and non-event day scores were calculated by month. Thus, each ADRS customer had two average performance scores for each month: an average Super Peak Period performance score for event days and an average non-event day peak period performance score.

The monthly performance scores for all event days each month were then averaged again into an overall score for each ADRS customer for the period of June 2004 through September 2005. An overall score averaging all the non-event day score by month was also calculated for each customer for this period.

ADRS participants were then sorted according to the overall scores. ADRS participants were first ordered by event day overall performance score. The list was then ordered by non-event day overall performance score. For example, if two ADRS participants both have an overall event day performance score of 3, but one participant has a non-event day overall performance score of 3 while the second has a score of 2.5, they are ordered so that the customer with a 3 score for both event and non-event days is placed higher on the list.

ADRS customers with average performance scores of 2.8 or greater for event days and 2.0 or greater for non-event day peak period initial load drop were selected as *super savers*. Customers with average Super Peak Period initial load drop performance score of 1.4 or less were selected as *program cruisers*. Customers who showed increasing performance from month to month in the average Super Peak Period initial load drop were designated as *improved performers.* Customers with 2 or more months with missing data were excluded from the selection process.

Estimating per household kW load impact

Measurement of ADRS load impact at an individual household level is problematic, because control homes cannot be matched with ADRS homes on a one to one basis. In all of the load impact analyses we've conducted in this report, RMI compared the average load of all control homes with that of ADRS homes for each time interval, by consumption stratum. Comparing an individual ADRS home with the average load of all control homes at a given time interval is not informative either, as we would be comparing an average control load that includes both large and small homes to one ADRS home that may have large or smaller loads than the average control load.

Given the data available, we decided to determine household level performance according to each home's immediate load drop at 2 p.m. compared to the period immediately prior, at 1:45 p.m. Furthermore, the 2 p.m. load drop would be scaled to the ratio of adjusted statewide average load reduction to the average 2 p.m. load drop. We judged this to be the best compromise to determining individual household performance, given the inability to compare against a control group at the individual household level. This “pre-curtailment” approach has been studied as an approach for automated demand response baseline calculations for individual customer accounts[5].

Thus for each ADRS customer, we began with the calculation of immediate load drop relative to 1:45 p.m. to 2:00 p.m. load on event and non-event days as described above, for each month from July 2004 through September 2005. The ADRS homes and their load impact results were segregated by high and low consumption strata.

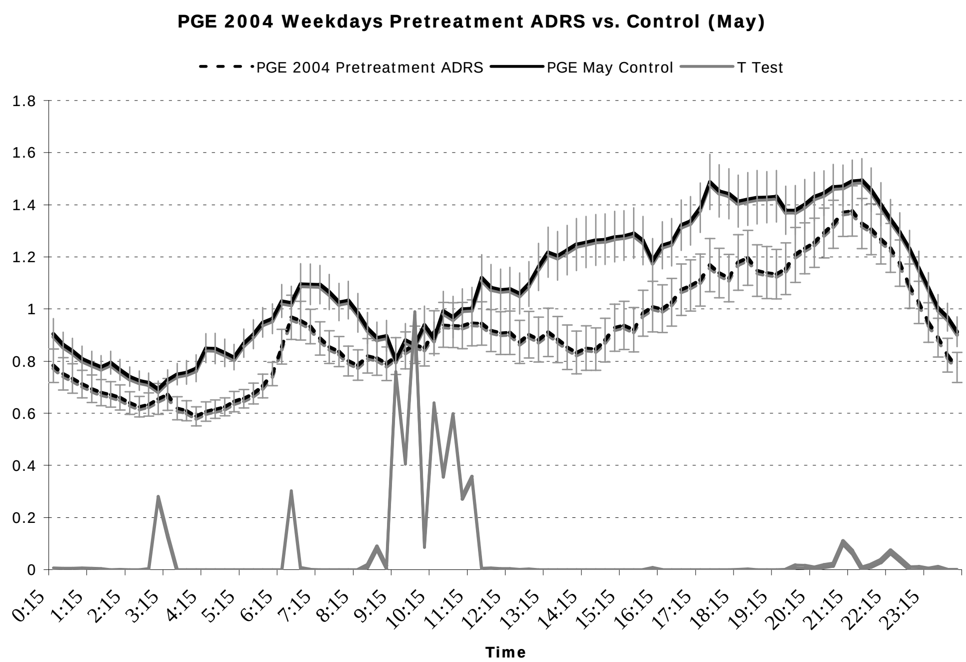
To scale the 2 p.m. load drop to the ratio of adjusted statewide average load drop, the average 2 p.m. load drop was then calculated for all homes combined, according to high and low consumption strata. Next, we calculated the ratio of the adjusted statewide average load reduction, which is based on the results of our 2004 and 2005 load impact evaluation, to the average 2 p.m. load drop of all ADRS homes. A separate ratio was calculated for each month from July 2004 through September 2005 for event and non-event days. This ratio was then multiplied by the immediate load drop at 2 p.m. for each individual household by consumption stratum. Once the adjusted immediate load drop at 2 p.m. for each household was calculated, RMI calculated the percentage of homes in each stratum whose load drop equaled or exceeded a given level on average.

[5] “Development of Uniform Protocols for Demand Response “Peak Savings” Calculations: A Review of Existing Methods and Recommendations for Uniform Protocols” Miriam L. Goldberg, CEC Staff Workshop, August 15, 2002

Appendix B
Statistical Tools

Figure 1 shows an example of a load curve from the pretreatment period, which RMI defined as the days in June 2004 prior to installation of technology in ADRS homes. The dashed line represents the average load by 15-minute time series for all the ADRS data in the pretreatment period. The solid black line represents the average load in the control group by fifteen-minute time series of all the days there are Pretreatment ADRS data. In a non-biased sample, one would expect to see the ADRS consumption curve closely matching that of the control (since it was assumed the ADRS group was biased). The issue is then how close would ADRS and control consumption have to be in order to consider them essentially the same. The statistical tools described below were used in the analyses to resolve this issue as well as many others in the qualifying and quantifying of bias in the ADRS sample.

Figure 1
PG&E Pretreatment



Confidence Intervals

Note the bars indicating the "confidence interval" around each data point on each line in Figure 1. Each point is an average of the data in our sample, but there were only a limited number of participants in the program.

This sample average is a proxy for a "population" average for whatever the sample represents. The confidence interval is a range around each average where the actual "population" average likely is (how "likely" is set by a predetermined interval). In the case of Figure 1, the 80% confidence interval is drawn around each data point, which translates into saying that there is an eighty percent chance that the actual population average for that time period is in this range.

Confidence intervals are also useful in giving some indication of how much variation or “noise” is in the data. Consider a city with an average temperature of 76^{o}F. The average of 76^{o}F could mean that day to day, the actual temperature hovers around 76^{o}F, or it could mean that the temperature varies from 105^{o}F to 55^{o}F. For the sake of the analysis, data with a lot of noise are more difficult to draw conclusions from, so information was included on the variability of the data in all of our analyses.

Student's t-Test

Another statistical test to assess whether differences between two samples are significant is called the t-test. The t-test is based on a family of distributions closely related to the normal distribution of probabilities that are influenced by sample size. While the Normal distribution is a good measure of distribution around statistics of interest (e.g. averages of load consumption) for large populations, t-distributions are typically used for sample sizes of less than 120 units. t-values along a t-distribution correspond to a range of probabilities (a.k.a. p-values) that a measured statistic (e.g. average load consumption) are *the same.* Thus, p-values resulting from the t-test range from zero to one. A p-value close to one indicates that two samples, such as the ADRS and control homes, have the same average loads for a particular fifteen-minute time interval. A p-value close to zero, then, means that average loads between ADRS and control homes are *different*. Note that throughout the day in Figure 1 the t-test is close to zero, indicating that the ADRS and control average load profiles are "significantly" different for most of the day.

The term "significant" refers to a result that is meaningful compared to one that is not. For instance, in all the analyses, a "significant difference" is defined as when the p-value for the t-test result is less than 0.05. This would mean that there is a five percent probability that two averages are the same. At or above 0.05, the averages are considered to be the same. Below 0.05, the averages are considered to be different.

Correlation

The correlation coefficient in statistics is a way to quantify the strength of a relationship between two independent variables. This coefficient falls between negative one and one: negative one indicates that the variables show a very strong *negative* correlation (as X increases, Y decreases) and one shows a strong *positive* correlation (as X increases, Y increases). A coefficient of zero indicates that there is no correlation at all. Another way to present this coefficient is in the form of R^2, which is simply the coefficient squared (so that it always falls between zero and one). This statistical tool was used to investigate how various variables affected the differences between ADRS and control groups.

Appendix C
Selection Bias Analysis

A07 Selection Bias and Load Impact Adjustment

In a January 16, 2004 draft program report**,** Charles River Associates (CRA) performed an initial investigation into the presence of selection bias between Statewide Pilot Program (SPP) participants on the critical peak pricing rate (CPP-F) and the control group in all three utility service territories across all climate zones. Subsequent to CRA's investigation, the California Energy Commission (CEC) conducted an independent evaluation of SPP load impact performance using a slightly different methodology for identifying and applying a selection bias adjustment. This section presents the two methodologies used in measuring the selection bias of SPP participants, and discusses RMI's rationale for choosing a selection bias adjustment approach that more resembles CEC's methodology than CRA's.

In 2003, PG&E, SCE, and SDG&E began recruiting residential and small commercial customers for the pricing-only Statewide Pilot Program. Eligible customers were invited to opt-in to the program. The three utilities sent out marketing materials, which included a pitch targeting working households not at home during the day. The pitch claimed that homeowners who did not already consume energy during the specific daytime hours when the program's experimental rate would be high would save money by participating in the program (see Figure 6). In essence, the program actively invited potential participants to join who would save money without any change in behavior. The program thus targeted potential free riders, and the SPP participants were found to have a bias towards lower peak period and overall consumption.

During the 2004 Pilot Program analysis, CRA, CEC, and RMI used different approaches in determining the use reductions due to the SPP and ADRS programs. CRA used a difference of differences approach based on regression models built from pre-treatment data, treatment data, and household surveys while CEC used a difference of differences calculated from the actual data. RMI made no adjustment for self-selection bias in the analysis of the A07 group in its December 2004 report[6] but is now restating the 2004 A07 results using the difference of differences method from actual data. The A07 homes used in the ADRS summer 2004 load impact evaluation are the subset of SPP participants in climate zone 3.

CRA examined the differences between the mean electric energy consumption for peak and off-peak usage in the pretreatment period[7]. A t-test was used to determine if the difference between the means was statistically significant. The differences in mean values during the peak hours in the pretreatment period were significant for climate zones 2, 3, and 4, indicating selection bias.

[6]Rocky Mountain Institute (RMI), *ADRS Load Impact Final Report*, December 28, 2004.
[7]CRA defined the pretreatment period as before a customer was put on the CPP-F rate. This definition problematic in that it is not clear the customers waited to be put on the rate to change their behavior. It is possible that treatment customers changed their behavior after receiving informational packets prior to be put on the rate.

Figure 6. Example welcome package material for SPP participants[8]

Example 3. No Air Conditioning: Customers At Work

CHART 9

Routine Electricity Cost

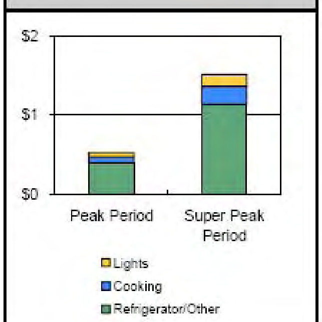
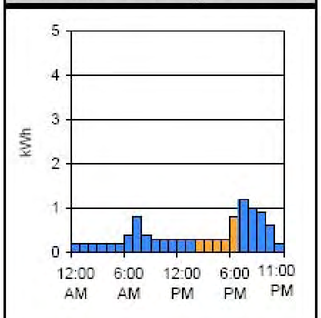


CHART 10

**Hot Summer Day
Routine Electricity Use = $2**



Weekday Electricity Use Habits

Patty and John both work away from home and usually return to their apartment between 6:30 p.m. and 7:30 p.m. They live on the coast and do not have air conditioning. Because of their work schedule, they use their dishwasher after 7 p.m. and do laundry on the weekends.

Patty and John use less electricity during the Peak period, Monday through Friday, 2 p.m. to 7 p.m., than the average customer. Since most of their electricity will be charged at the new, lower Off-Peak rate, their bill is likely to go down without shifting or reducing their electricity use.

> **Chart 9: Electricity use between 2 p.m. to 7 p.m. on Super Peak Days is about three times more expensive than on other weekdays. However, since Patty and John are not home in the afternoon, their Super Peak costs are under a dollar more than on other weekday afternoons.**

On most days, the cost of their electricity use is between about $1 and $2. Because they are not home during most of the Peak period and don't have air conditioning, Patty and John's cost on a Super Peak Day is about $2. As shown in Chart 9, they do not have many options to shift or reduce their use during a Super Peak period because most of their electricity use is for appliances that are always on, like their refrigerator.

Cost Saving Ideas: Shift & Save

However, even Patty and John can lower their monthly bills. In areas where lights are usually on in the evening, compact fluorescent lamps will pay for themselves with electricity cost savings in about a year. This is an investment that will be easy to take with them if Patty and John move in the future.

Chart 10 shows Patty and John's hourly electricity use on a Super Peak Day, assuming that they take no action to shift or reduce their use. Shifting what they can control after 6 p.m. will save them under $0.20 that day.

29

8 SHIFT & SAVE PRICING PLAN

---

[8]Charles River Associates (CRA), *Statewide Pricing Pilot Summer 2003 Impact Analysis, Appendices*. August 29, 2004. Appendix 4, page 29.

CRA's regression model and difference of differences approach

In order to correct for the apparent bias, CRA performed a "difference of differences" analysis to subtract the pretreatment difference between the participant and control groups from the difference observed after the program went into effect. However, they did not simply compare the measured differences during the pretreatment period to those measured after the SPP began in order to arrive at a measure of actual load reduction, minus selection bias. Instead, they performed a regression analysis. CRA stated that simple comparisons of means can be misleading because they ignore the influence of various other variables that may also affect energy usage, such as weather, appliance holdings, socio-demographic factors, income, and attitudes about the environment and the utility. Thus, they employed a multivariate linear regression model to control for the difference in weather between pre- and post-treatment periods and for other variables between the control and treatment groups and the population at large.

Steps Taken to Arrive at the Final Measure of Bias through the Regression Model[9]

1. Separate regression models were estimated for peak usage, off-peak usage, and daily usage for each participant by rate treatment and climate zone in the *pretreatment period*. All data from each customer were used to create a model with the following characteristics:
    - Model Form:
      kWh = constant + coefficient_1* variable_1 + coefficient_2* variable_2 + …

    - There were 22 variable terms in total, some of which included interactions between the variables.
    - Several variables were binary terms, indicating that they could take only two values. This was represented with a 1 or a 0.
2. The models were then used to predict kWh, the dependent variable, for all participant and control populations in the pretreatment period. The average values for each variable were calculated within each target population and were multiplied by the regression coefficients solved for in step 1 to predict kWh. SPP Pilot period weather, measured by cooling degree hours, was also used.
3. The kWh difference between control and participant customers was calculated, providing a pretreatment difference in consumption for the two groups.
4. Steps 1 through 3 were then repeated for SPP *pilot period* data and an estimated difference in consumption between control and treatment populations was calculated.
5. The "difference of differences" was then calculated in order to attain an <u>unbiased estimate of the treatment impact after adjusting for differences between the groups</u>.

A summary of the procedure is displayed as follows:

- Δ_1 = Participant Predicted kWh – Control Predicted kWh (in the pretreatment period)
- Δ_2 = Participant Predicted kWh – Control Predicted kWh (in the treatment period)
- $\Delta = \Delta_2 - \Delta_1$

[9] All variables in the regression model are displayed on p.67 of the CRA January Draft Report.

*For super peak day analysis, CRA used the data from their 12 high system load days in May and June in the pretreatment peak regression model.

CEC's difference of differences approach

The CEC used the differences of differences approach using actual data. This entails taking the difference in consumption between treatment and control groups during peak and off-peak periods in the pretreatment period and subtracting that from the difference in consumption between the treatment and control groups during the peak and off-peak periods in the treatment period. This accounts for possible selection bias in the treatment group. In contrast to CRA, the CEC defined the pretreatment period as the month of June 2003. Though an uncomplicated approach, this definition is problematic in that it assumes that all customers began modifying their behavior after June. In reality, many customers were not aware of when exactly they were placed on the experimental CPP rate, and could have begun modifying their behavior as soon as they agreed to join the program.

Both CRA and CEC defined a proxy for super peak days as the 12 maximum system load days in the months of May and June. For climate zone 3 participants, the differences in peak period energy use were significantly greater on these hotter days than the cooler ones. For example, participant (A07) peak period use was 20 percent less than control customer use. Total daily energy use was 13 percent less for participant customers than for control customers in climate zone 3 on these hot days. No significant difference in daily energy use was observed in climate zone 3 on the cooler pretreatment days (CRA January 2004 Draft Report, p. 53).

Table 6. Comparison of CRA and CEC's summer 2003 SPP results for climate zone 3 participants

CPP Days Peak Period Comparison (Climate Zone 3)

	Zone 3 CPP Change (kWh/h)	Savings (kWh)	% Reduction
CRA†	0.22	1.1*	13.37%
CEC††	0.30	1.5	16%

Non-event Weekday Peak Period Comparison (Climate Zone 3)

	Zone 3 CPP Change (kWh/h)	Savings (kWh)	% Reduction
CRA†	0.08	0.4*	5.59%
CEC††	0.11	0.6	8.5%

CPP Days Peak Period Comparison (Statewide)

	Zone 3 CPP Change (kWh/h)	Savings (kWh)	% Reduction
CRA[1]	0.15	0.75*	12.50%
CEC††	n.a.	n.a.	12%

* Value calculated by author
†: Charles River Associates (CRA), *Statewide Pricing Pilot Summer 2003 Impact Analysis* Final Report October 11, 2004. p. 7
††: Pat McAuliffe and Arthur Rosenfeld, *Response of Residential Customers to Critical Peak Pricing and Time-of-Use Rates During 2003 and 2004*. January 17, 2005. September 23, 2004. p. 4

Table 6 summarizes the results of the CPP programs based on CRA's and CEC's independent analyses. Note that the CEC and CRA are evaluating the success of the CPP-F program for June through October of 2003, while the RMI control-A07 comparison looks at performance from June through September of 2004.

Using a difference of differences approach, the CEC calculated a statewide reduction due to the CPP-F rate of 12%. Using a multivariate regression based model CRA calculated a 12.5% reduction in energy use due to the CPP-F rate. The 0.5% difference between the two approaches is negligible, especially when uncertainties in the CRA modeling process demand values are reported to no more than 2 significant figures.

RMI's difference of differences approach

Our restating of the 2004 ADRS load impact results in volume 2 of this report includes a selection bias adjustment for A07 homes that is similar to the CEC's difference of differences approach. The additional benefit of using multivariate regression based models approach that CRA adopted is not great in this case (see Table 6). The regression model approach is strong because it tries to control for weather differences between pre- and post-treatment periods. This is necessary because the temperature in the pretreatment months was generally lower than in the summer when critical peak days would be called. Thus, cooling degree hours and the binary central air conditioning variable appear to be a useful control in arriving at a true measure of bias. They allow for treatment period weather conditions to be used when analyzing pretreatment data and they control for whether or not a customer has central air conditioning to respond to the weather.

However, the additional variables regarding household characteristics have questionable basis in the CRA model. These variables were introduced to try to control for differences between homes in the sample and the population at large. However, in CRA's final report, published in October 2004, an analysis of summer average daily use from 2002 from the three investor-owned utilities indicated that there was no bias in the treatment and control samples compared with the population as a whole. These results question the need to include additional household characteristic variables that were aimed at correcting for differences between the control and treatment groups and the population at large.

There is also no evidence of any substantial model selection analysis that quantitatively justifies the model choice. There is no presentation of any statistics that compare model explanatory power to a loss of precision with the addition of new terms. The only justification for the large regression model is that all "observable" factors need to be accounted for[10]. However, some variables in the selected model may have no influence or may be redundant and/or unnecessary. There is no evidence of any quantitative comparison that would support the decision to use the selected model over an alternative with fewer terms. In addition, there is no attempt to state that model selection statistics are unnecessary through some type of qualitative justification of why every term is relevant to the research question at hand.

Figure 7 shows the selection bias adjustments that we applied to the statewide A07 loads. These differences were calculated based on the control and A07 groups' loads during June 2003, which the CEC designated as the pre-treatment period. May 2003 data were excluded from pretreatment

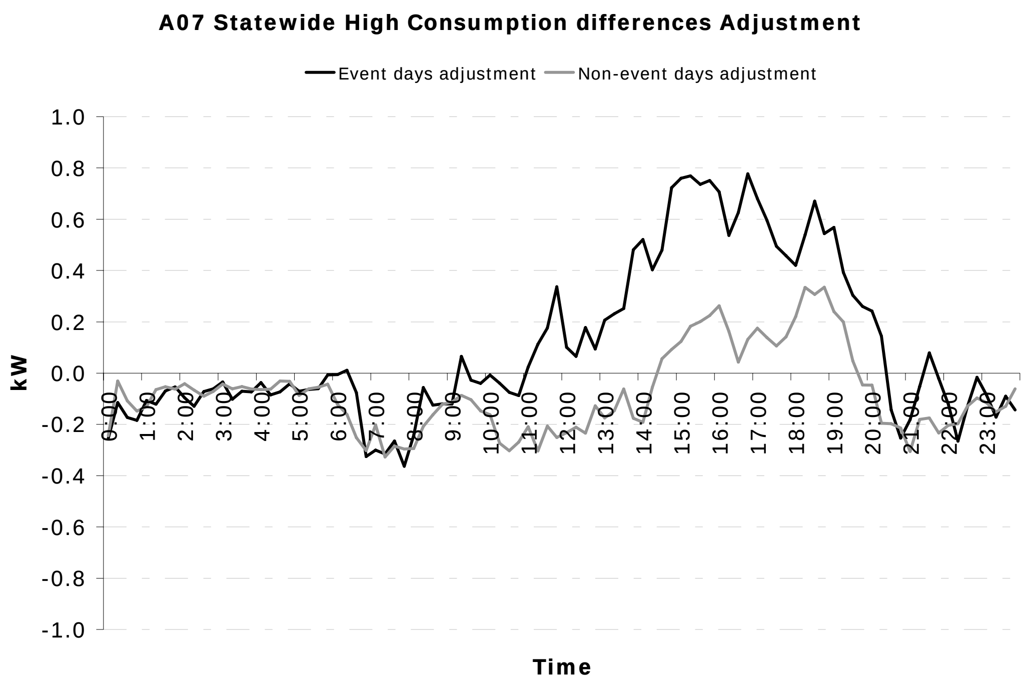
[10] Charles River Associates (CRA), *Statewide Pricing Pilot Summer 2003 Impact Analysis, Draft Report*. January 16, 2004. p. 66

analysis due to its scarcity. For the same reason, bias adjustment for A07 was not evaluated on a utility-by-utility basis from June 2003 data due to scarcity.

From the pre-treatment data furthermore, we extracted seven of the hottest days during the pre-treatment period to simulate event days during the pilot period. These were also the same days used by the CEC in its original selection bias adjustment for A07[11].

The rest of the June 2003 data were used to simulate the non-event days during the pilot period. We calculated hourly loads averaged over all houses for the control and A07 groups in the CPP and non-CPP subsets and then calculated the difference between these averaged loads (Control-A07).

Figure 7. Selection bias adjustment applied to A07 load impact results for the 2004 ADRS pilot



The plot of the differences adjustments is as expected. Statewide, the A07 group consumed less load than the control group during the peak period (2pm-7pm) on both hot and cool days. As such, the differences adjustment was used to shift the A07 loads higher to match the control customer consumption during these hours. The differences during off-peak hours are very small—less than 0.05 kW. The peak period differences are much greater on the simulated event days than the simulated non-event days. On simulated event days, control-A07 differences averaged 0.61 kW across the peak period. On simulated non-event days, control-A07 differences averaged 0.14 kW across the peak period.

[11]June 2, 16,17, 25, 26, 27, 30

ADRS Selection Bias and Load Impact Adjustment

In 2004, RMI's load impact analysis method was a straightforward engineering approach where actual meter data for ADRS participants were compared against two sample populations on Super Peak and non-Super Peak days. The first sample consisted of single-family homes with central air-conditioning in climate zone 3 with on standard tiered electricity rates. This group is named the "A03" control group. In addition, a second sample was used that was a subset of the single-family homes with central air-conditioning in climate zone 3 on an experimental, dynamic critical peak pricing rate (CPP-F). These customers were also participants in California's Statewide Pricing Pilot program (SPP), and are named the "A07" homes. The simple differences approach in analyzing ADRS load impact against the two sample populations ignored any pre-existing differences between the three groups.

The ADRS pilot was structured such that homeowners opted into the program in the manner of a typical utility demand-side management program. It is possible that participants who self-selected into the ADRS pilot possessed non-random characteristics that differ from the general population at large. Based on these discussions during the summer 2004, RMI was tasked to assess the presence of bias in the ADRS home selection and adjusted the load data accordingly.

Methodology overview

In an ideal experiment, relevant data are collected on a population prior to the instigation of the experimental treatment—in this case, ADRS participation. The "pretreatment" data are thus used to confirm that a population is truly random and representative of the general population at large, or to measure any differences between a selected population and the population at large. In the case of ADRS, ideal pretreatment data entails the installation of interval meters on a population of homes *one year* before the actual initiation of the pilot program, such that fifteen-minute interval data would have been collected the summer period before pilot activities began.

In reality, interval meters were installed on homes after participants opted to participate in the program. This not only limits the quantity of data available that can be considered pretreatment but also reduces the certainty that the data truly reflects "pretreatment" behavior. In light of the scarcity of true pretreatment data RMI employed additional data sets used as pretreatment proxies in the assessment of potential bias in the ADRS participant population. Each dataset has inherent weaknesses because it is technically just a proxy for true pretreatment data. However, the use of several data sources to create a composite picture of ADRS consumption behavior compared to control creates a more complete picture of potential bias than any of the methods alone.

Ultimately, one qualitative and three quantitative analyses were conducted using the following data sets:

1. ADRS program recruiting and welcome materials,
2. Interval meter data during ADRS "pretreatment" weekdays prior to ADRS (GoodWatts) technology installation,
3. Interval meter data during 2004 summer weekends, and
4. Monthly kWh billing meter data from summer 2003

First, recruitment and welcome materials for the ADRS pilot program were reviewed to determine whether marketing messages were framed such that they would encourage homeowners already consciously conserving loads to volunteer for the program. While the SPP program explicitly invited homeowners already conserving energy to join, the ADRS pilot marketing messages focused on the use of technology to enable energy and cost savings and did not try to attract customers based on lifestyle or ethos. Furthermore, the utilities specifically wanted to downplay the potential savings aspect of the program since they wanted to ensure that customers did not expect savings, but rather, would realize savings only if they changed their behavior during peak and super peak periods. Because customer recruitment targeting was limited to physical household parameters such as single-family homes with central air conditioning, RMI's initial hypothesis is that customers who were recruited into the ADRS program were not much different in their consumption behavior on average from the general population at large.

Second, ADRS pretreatment data were evaluated to detect differences, if any, with control homes before the ADRS program. The pretreatment period was defined as non-holiday weekdays prior to installation date of ADRS technology. This parameter was selected because installers were required to enter participant homes in order to install the technology. Interval meters were installed outside the home, without the need to contact customers for scheduling. RMI thus hypothesized that ADRS participants believed they were placed on the critical peak pricing rate (also CPP-F) at the time of technology installation and thus began to change behavior at that time[12].

A major weakness in the pretreatment data is that they were collected *after* residents volunteered to participate in the pilot. It is also debatable whether or not "treatment" technically began when residents received marketing and recruiting. Some customers reported to Invensys Climate Controls that they believed that they were placed on the CPP-F at the time of the interval meter installation, which in most cases, occurred prior to the installation of the GoodWatts technology. Thus, these customers might have begun modifying behavior even before GoodWatts technology installation. Using the meter install date as the cutoff for pretreatment data, however, resulted in only a few day's of data from only a few homes. This would have eliminated the ability to use "pretreatment" data at all.

For lack of a more robust source of pretreatment data, the pre-Goodwatts installation data were considered for this analysis. Another weakness in using pre-GoodWatts technology installation as a proxy for pretreatment data is that the technology installation dates were only available for two utilities, PG&E and SCE. Thus, bias selection analysis could only be conducted for PG&E and SCE, but not for SDG&E ADRS customers.

Third, June – September 2004 weekend and holiday data for both ADRS participants and augmented control customers were consolidated and compared. The fundamental weakness of weekends and holidays data were that weekend consumption behavior was fundamentally different from weekday consumption behavior given differences in occupancy patterns. For this reason peak and Super Peak pricing did not occur during weekends and these data served as an imperfect approximation of behavior during peak and Super Peak hours. Furthermore, participants were in the program during summer 2004 weekends and had likely already modified their behaviors. However, the quantity of

[12]In reality, ADRS participants were put on the CPP-F rate on the billing date following technology installation, but most participants did not know when that was.

data available for this period was greater than that during the pretreatment period, providing a more statistically robust picture of off-peak behavior. Weekends were analyzed because participants were subject to a relatively low, non-dynamic off-peak rate that facilitates more “normal” behavior, with less influence of the ADRS program or peak/Super Peak electric rates.

Finally, monthly billing data for January – December 2003 were compared, for the ADRS and augmented control customers. The weakness of monthly consumption data was the inability to interpret behavior at finer than monthly time scales, as compared to intraday behavior provided by interval data. This data set represented an additional benchmark against which participant and control customer behavior could be compared.

Given that each of the three quantitative proxies for pretreatment data have their weaknesses, conclusions about bias *cannot* be made based on interpretation of each data source by itself. Rather, the determination of ADRS bias must rely on the simultaneous evaluation of all data sets together, based on the composite results produced.

Overview

A qualitative review of recruiting and welcome package materials confirmed there were no marketing messages specifically targeting homes that were already conserving energy. Monthly billing, pre-ADRS technology data (pretreatment) and summer 2004 weekends analyses produced consistent patterns in the orientation of ADRS customers’ consumption relative to the augmented control sample. While the 2003 monthly billing analysis revealed that these consumption differences were not statistically insignificant, more detailed information available through pre-ADRS technology installation and 2004 summer weekend interval load data revealed significant differences in ADRS-control consumption, particularly during the peak period hours of 2 p.m. to 7 p.m. Consumption differences outside of peak hours were mixed, with the majority of differences statistically insignificant.

The orientations, furthermore, differed by utility. For PG&E and SDG&E, ADRS customers consistently had **lower** load than their associated customers in the control sample. On the other hand, SCE ADRS customers consistently had **higher** load than their associated control sample customers.

The pretreatment dataset was highly problematic, primarily because it was extremely thin. SDG&E customers could not be evaluated, as ADRS technology installation dates were not available. The number of PG&E and SCE pretreatment homes by utility declined significantly beyond May 15, 2004. Homes in which ADRS technology was installed before other customers had more pretreatment data recorded than other homes. Other homes in which ADRS technology was installed later in May did not have load data available earlier in the month for analysis. This uneven weighting of homes in the pretreatment sample during May 2004 introduced significant noise in the pretreatment data analysis.

In an attempt to extract a meaningful dataset from the pretreatment dataset, we analyzed them from a variety of perspectives: by temperature, by geography, by utility, and finally by consumption stratum. Neither temperature nor geography yielded consistent results that could be explained by the available data. However, segmenting the pretreatment data by utility and by consumption stratum provided load curves resembling those of load curves during the ADRS pilot period. The signals that emerged from the pretreatment data, albeit weak, showed statistically significant differences between ADRS and control between 2 p.m. and 7 p.m.

Segmentation of weekends load data yielded consistent results with the pretreatment. Furthermore the ADRS vs. control differences measured using the two datasets were statistically similar. Combining both summer 2004 pretreatment data and weekends data into an average bias adjustment is not correct, as such a combination actually introduces more error than using one or the other data sources alone. The question then is which data set should ultimately be chosen for use in the bias adjustment. Given that the data set for summer 2004 weekends was larger and therefore more robust, it was selected for use in quantifying the difference adjustments for all load impact analyses.

ADRS program recruitment and welcome materials review

The first assessment of ADRS selection bias was a review of the recruitment materials sent to all potential ADRS homeowners and the welcome materials sent to those who later enrolled in the program. RMI received one non-utility specific version of these materials and assumed that homeowners received identical or nearly identical versions. We searched the materials for passages that attracted participants based on lifestyle or ethos.

Both the recruitment letter (Figure 8) and welcome package were found to have consistent marketing messages. The ADRS program was marketed as a technological and cost-neutral way to reduce electricity use without the imposition of conservation on the homeowner. Specifically, the program was an opportunity for the homeowner to “test the latest in home energy management technology” and allowed the homeowner to “take advantage of a new electric rate” and potentially “save money depending on when [they] use [their] appliances.” There did not appear to be any direct marketing pitch to potential “free riders” – homeowners that already actively conserved electricity or whose behavior or occupancy patterns resulted in low on-peak usage.

In the recruiting materials, there is an appeal for "By reducing your electricity use during the 2 p.m. - 7 p.m. period on super peak days, you can avoid these higher prices, and also help reduce the demand on the energy system," and "The new rate also includes higher prices on 12 "Super Peak Days" when electricity demand is highest, and when saving energy can help avoid rotating outages." The "help reduce demand on the energy system" is an appeal to consumers who are conscious of broader problems on the grid. However, given the California electricity crisis, this would be just about everyone in the state. It is not a direct marketing pitch to potential “free riders” –homeowners that already actively conserved electricity[13].

[13]This observation is also consistent with results of Boice Duham Group’s market research conducted for the summer 2004 pilot, that most ADRS participants were not conservers prior to GoodWatts.

Figure 8
ADRS Recruitment Letter

March 1, 2004

Mr. A. Customer
1256 Caliente Lane
San Diego, CA 10011

Dear Mr. Customer,

Managing your home's energy use and saving on energy costs is important to everyone. Within a few days, you'll receive our invitation to participate in a new Pacific Gas & Electric program. You can also sign up early by responding to this letter. Our GoodWatts program, endorsed by the California Public Utilities Commission, gives you an opportunity to test the latest in home energy management technology.

Along with the equipment, you'll also have the opportunity to take advantage of a new electric rate that changes during the day, and allows you to save money depending on when you use your appliances.

By participating in our GoodWatts program you'll receive a new Internet-programmable thermostat for your central air conditioning, and Internet energy management tools, both at no charge. With these tools, you'll be able to manage your home's energy use from anywhere you can access the Internet. You can track your energy use hour-by-hour, and reset your air conditioner or swimming pool pump to maximize your savings.

Your invitation will include a complete description of the GoodWatts program, an application and a postage paid return envelope. This program is free and will be available to a limited number of customers with central air conditioning.

If you join and later decide the program isn't right for you, you can opt-out at ***any time***.

If you'd like more information about this program, please call 877-811-8700, or go to www.goodwatts.com. You can see how the technology works and get more information on your savings potential. You'll also be able to find out more about the electric rate that is part of this program.

Thank you for your consideration and for helping to shape the future of energy technology in California.

Sincerely,

Tim Vahlstrom
Principal Project Manager, Energy Program Services

P.S. By participating in this program, you'll receive appreciation payments of up to $100.

Some homeowners might have been attracted to the program because they saw the Goodwatts technology as a way to augment their energy conservation efforts. Other homeowners with above average energy use might have seen it as a way to save money while maintaining their consumption habits. The marketing materials did not implicitly attract homeowners with any specific consumption pattern. Thus, no conclusions could be drawn about how marketing and recruitment tactics induced bias on the ADRS sample.

Pre-ADRS technology installation analysis (pretreatment)

PG&E, SCE, and SDG&E provided fifteen-minute interval data for May 2004 for both ADRS and (augmented) control homes. A list of the dates on which ADRS technology was installed in participant homes was also requested and provided by Invensys. PG&E and SCE ADRS homes had these dates recorded but SDG&E dates were missing, so no pretreatment analysis could be performed for SDG&E ADRS homes.

For PG&E and SCE, ADRS customers were included in the calculation of daily average loads *until* the technology was installed. After the technology install date, ADRS homes were considered *actively* participating in the program and removed from the pretreatment set. Load curves shown in Figure 9 and Figure 10 for PG&E and SCE respectively were created by averaging all weekdays in the pretreatment period for the ADRS and control homes by utility. P-values for each time interval using two-tailed t-test are also plotted below both curves. Statistical tools used in ADRS bias analysis are described in APPENDIX B.

Figure 9
Comparison of ADRS vs. Control (Augmented) Consumption During Pretreatment Period, PG&E

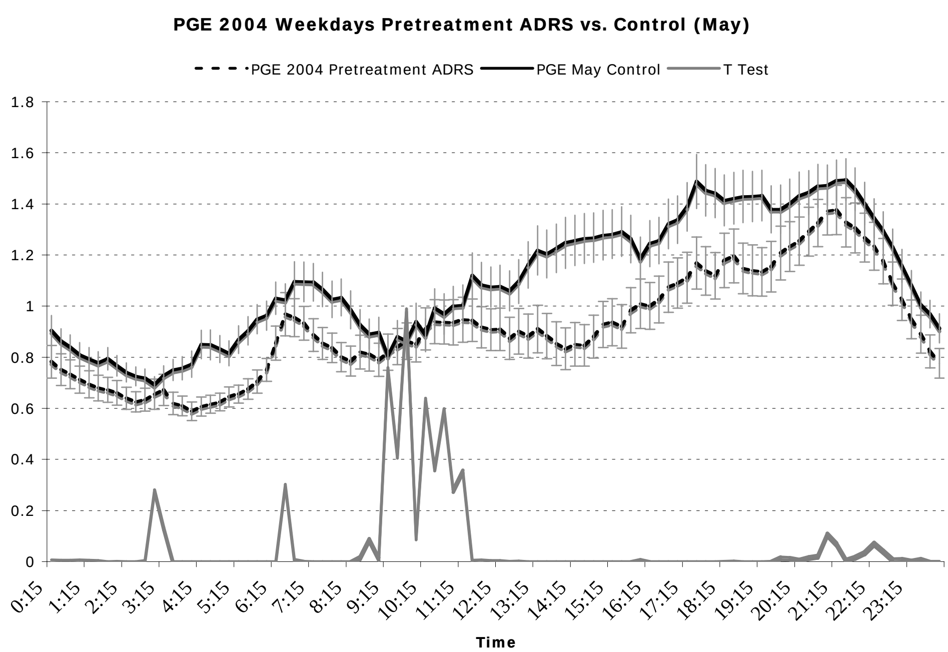
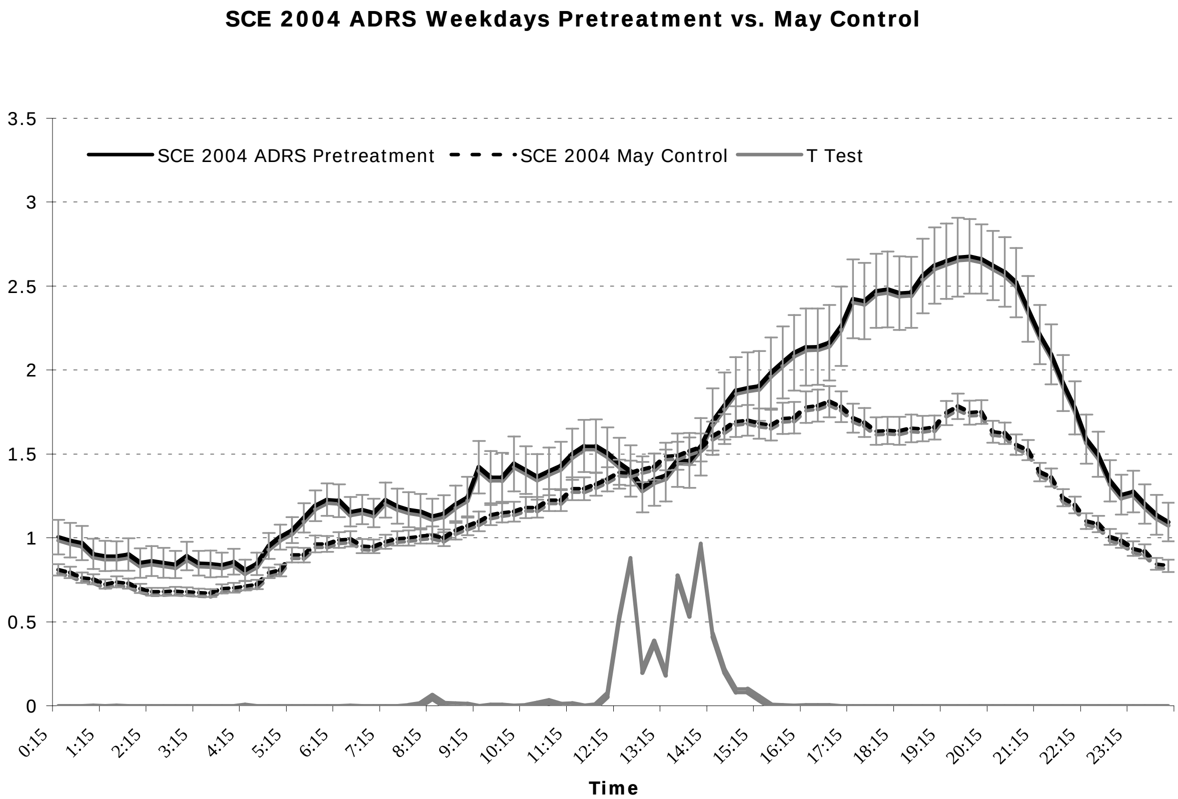


Figure 10. Comparison of ADRS vs. Control (Augmented) Consumption During Pretreatment Period, SCE



PG&E ADRS homes consumed less load than control homes by approximately 0.25 kW on average across the day. On the other hand, SCE ADRS homes consumed more load than the control homes by approximately 1 kW. This indicated that it would be inappropriate to apply the same differences correction across all utilities. The p-value generated using two-tailed t-test was close to zero throughout most of the average pretreatment day in both PG&E and SCE daily load profiles, indicating that differences between ADRS and control groups are statistically significant.

Table 7. Number of homes in ADRS and Control Groups Strata

	PG&E		SCE		SDG&E	
	ADRS	control	ADRS	control	ADRS	control
Low Consumption	24	2	4	14	15	3
High consumption	51	30	72	52	7	22
S2/S1	2.1	15	18	3.7	0.5	7.3

However, the size of the confidence intervals around each load value (at times nearly 0.5 kW) suggests that there may be ways of dividing the pretreatment data which would result in cleaner averages with less noise for use in a differences correction. Table 7 shows the number of homes in each stratum for each utility as well as the ratio between high and low consumption strata for each group. These differences in the relative numbers of ADRS versus control homes by stratum distorted the average load profiles of the ADRS versus control pretreatment data set as a whole (Figure 9 and Figure 10) and added noise as well.

A more detailed examination of the variations underlying the average load curves calculated for each utility was performed.

Table **8** shows the number of houses in the pretreatment data set for each weekday during the month of May. The count reveals that pretreatment data became progressively thinner later in the month as more and more houses were fitted with Goodwatts. A count of the number of pretreatment days each ADRS home had in the pretreatment data set is shown in Table 9. This analysis elicited another interesting result: each ADRS home *did not* have the same number of pretreatment days in the data set. Thus, each ADRS home was, in effect, weighted differently in the pretreatment data set. This uneven weighting of homes in the pretreatment sample during May 2004 introduced significant noise in the pretreatment data analysis. We next examined the data on a daily basis in to better understand these temporal variabilities.

Table 8. Count of Pretreatment Homes Through May

PG&E All Homes				**SCE All Homes**		
Date	ADRS	control		Date	ADRS	control
5/3/04	40	14		5/3/04	30	65
5/4/04	40	32		5/4/04	29	65
5/5/04	40	32		5/5/04	27	66
5/6/04	40	32		5/6/04	24	66
5/7/04	40	32		5/7/04	21	66
5/10/04	39	14		5/10/04	19	66
5/11/04	37	32		5/11/04	19	65
5/12/04	34	32		5/12/04	20	66
5/13/04	31	32		5/13/04	20	65
5/14/04	28	33		5/14/04	22	65
5/17/04	21	14		5/17/04	20	66
5/18/04	18	33		5/18/04	18	65
5/19/04	14	33		5/19/04	17	66
5/20/04	13	33		5/20/04	18	66
5/21/04	10	33		5/21/04	15	66
5/24/04	8	14		5/24/04	13	65
5/25/04	6	33		5/25/04	12	66
5/26/04	6	32		5/26/04	8	66
5/27/04	6	32		5/27/04	6	65
5/28/04	5	32		5/28/04	5	65

Table 9. Pretreatment Days Count by Home

PG&E ADRS				SCE ADRS	
Home ID	# Pretreatment days	Home ID	# Pretreatment days	Home ID	# Pretreatment days
D0056A	31	D0092A	11	1923063	31
D0063A	31	D0037A	10	2781874	31
D0078A	31	D0043A	10	1488196	28
D0086A	31	D0077A	9	1338434	27
D0105A	27	D0106A	7	2953927	26
D0062A	25	D0093A	5	6420672	26
D0060A	24	D0104A	4	3721380	25
D0029A	23	D0111A	1	4990698	25
D0041A	23	D0112A	1	6945018	25
D0022A	20			7025772	24
D0076A	20			6717363	21
D0094A	20			6639809	20
D0101A	20			2004217	19
D0036A	19			2396864	19
D0069A	18			1417171	18
D0074A	18			202195	17
D0080A	18			5517643	17
D0099A	18			1944323	15
D0049A	17			2669906	14
D0059A	17			4115019	14
D0072A	17			4442645	14
D0042A	16			4877233	13
D0045A	16			6604723	8
D0088A	16			2573075	7
D0067A	14			5502689	7
D0070A	14			856949	6
D0073A	14			4267602	6
D0084A	14			4474368	6
D0032A	13			6969859	6
D0075A	13			210547	5
D0110A	13			1148207	5
D0091A	12			1842240	5
D0100A	12			729204	4
D0107A	12			2256300	4
D0047A	11			6670610	4
D0058A	11			2724750	3
				2985786	3
				1344171	2

Figure 11 and Figure 12 show six daily charts from the pre-GoodWatts installation period for PG&E and SCE. Each chart displays average load consumption for ADRS and control homes in 15-minute intervals. The t-test p-value and difference between control and ADRS are also plotted for each corresponding interval. Eighty percent confidence intervals are drawn as bars around each data point. Average peak temperature and sample size are indicated on the legend for both ADRS and control samples.

Figure 11. PG&E Pretreatment Days

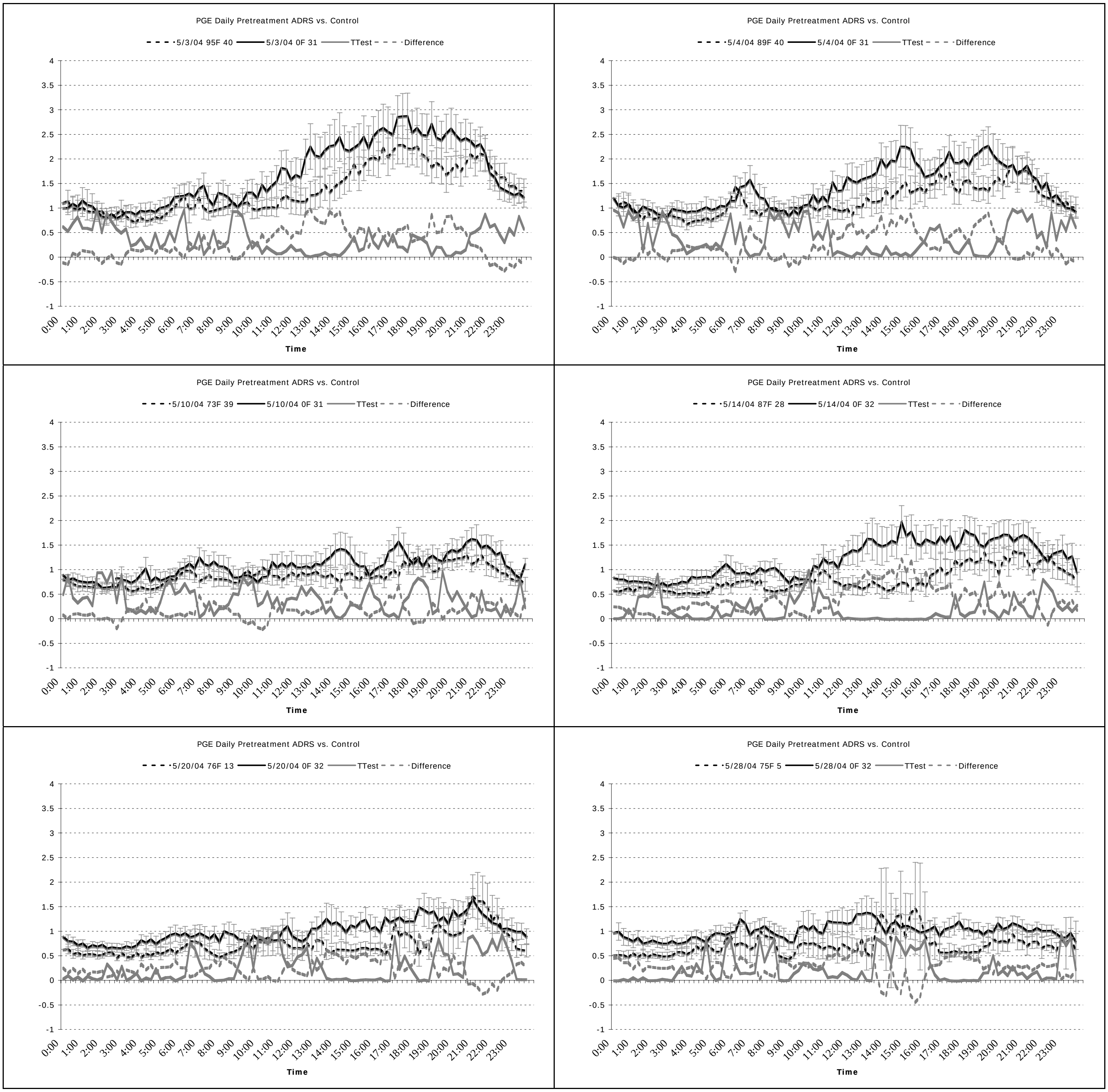
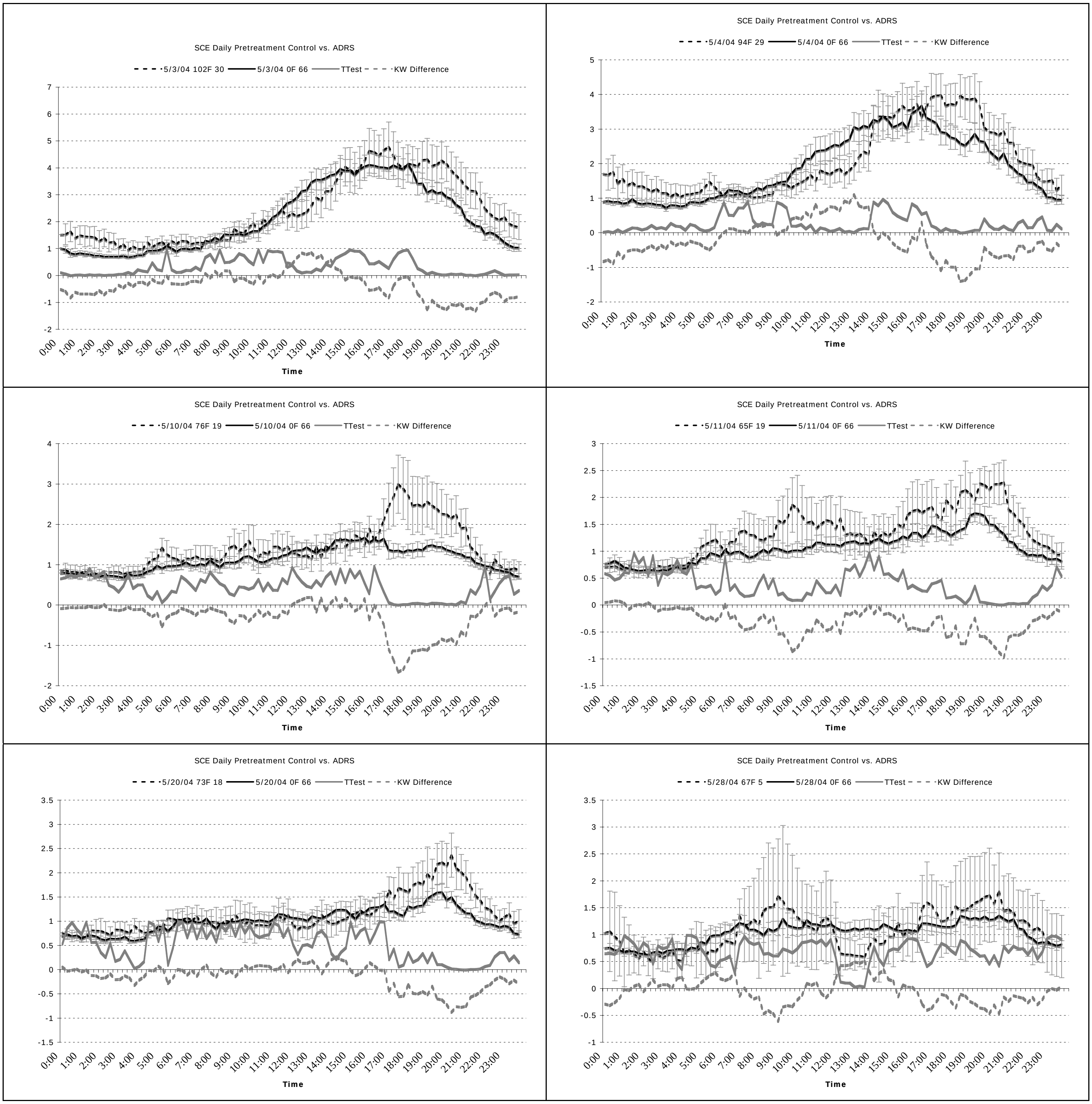


Figure 12. SCE Pretreatment Days



Looking across the individual days the ADRS and control curves did not seem significantly different based upon the T-test value and wide confidence intervals. T-test p-values frequently exceeded 0.05 and the confidence intervals between curves often overlapped. This observation was true for both PG&E and SCE. Each chart also displayed a lot of noise in the averages. The average daily load curves were not smooth but highly variable in an apparently random way. Furthermore, the shape of the load curves did not seem to follow any consistent pattern throughout the day. These characteristics suggested the need to segment the data to uncover a clearer signal in load consumption behavior.

)Temperature

The first attempt at segmenting the pretreatment data looked at whether differences in pretreatment consumption varied with temperature. Comparing load profiles from hot and cold days, temperature appeared to be a dominant exogenous factor in determining load. For example, compare ADRS and control loads on May 3rd and May 10th for PG&E shown in Figure 11 (95°F and 73°F, respectively) and May 3rd and May 11th for SCE shown in Figure 12. (102°F and 65°F, respectively). The hot day of May 3rd exhibited a more pronounced increase in the load difference between ADRS customers and the control group, from minimal differences in early morning to a maximum load difference of 0.9 kW by 1 p.m. (SCE) and 2 p.m. (PG&E). The cool days in May exhibited less load differences between ADRS and control customers, from minimal differences in the morning to 0.7 kW difference by 2:30 p.m. (PG&E on May 10th) and 0.5 kW difference by 5:00 p.m. (SCE on May 11th). We thus hypothesized that the largest differences between control and ADRS loads also happened on the hottest days.

The maximum ADRS and control load differences vs. the average maximum temperature were plotted by for every weekday during the pretreatment period. If RMI's hypothesis about greater control-ADRS consumption differences on hotter days were true, we would expect to see a strong correlation between the two variables. Figure 13 and Figure 14 show these scatter plots for PG&E and SCE, respectively. The correlation coefficient for the chart was nearly zero, indicating that there was *no* correlation in average peak temperature and maximum kW difference for any utility. Thus the hypothesis about temperature and load difference was incorrect, and we next examined the possible relationship between load and geographical location of ADRS customers to discover a clearer signal in the pretreatment data.

Figure 13. PG&E Temperature Correlation

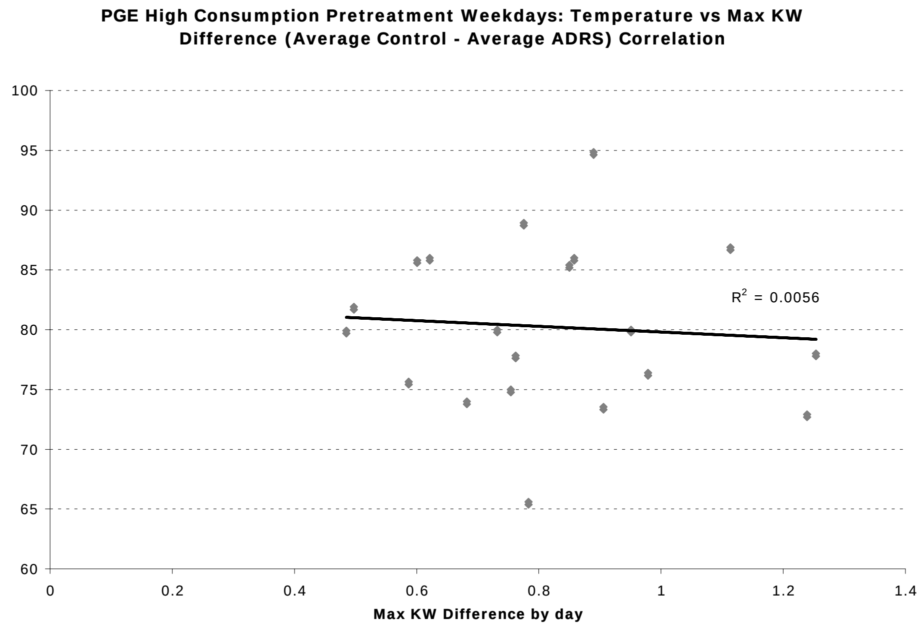
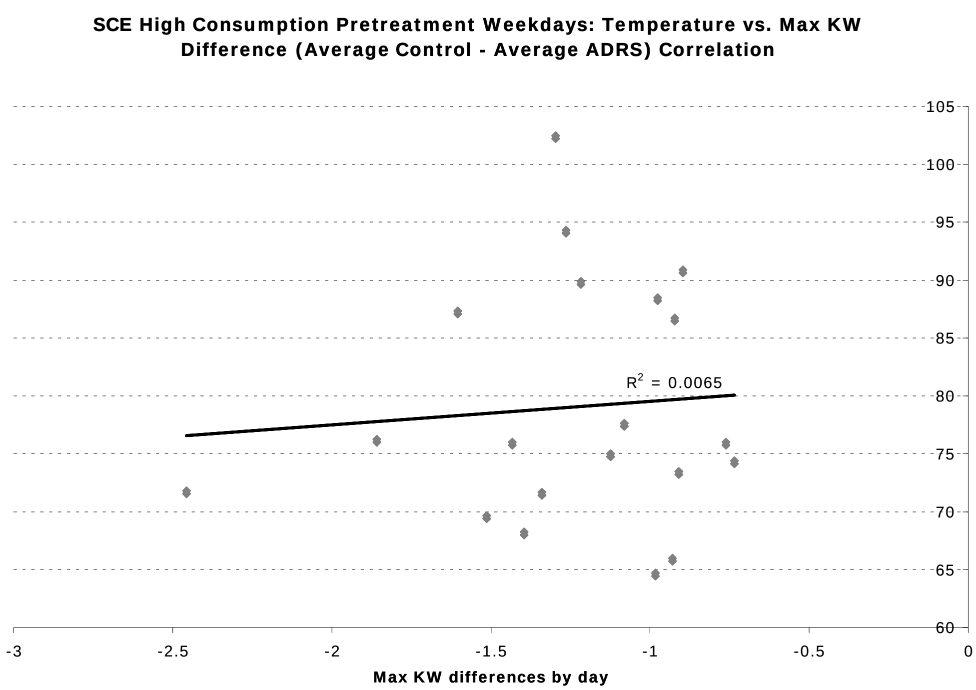


Figure 14. SCE Temperature Correlation



)Geography

Given the lack of results in the temperature study, we researched the geographic locations of ADRS homes to study whether they were possibly affecting the temperature analysis and load bias. Figure 15 maps of the locations of all the ADRS homes for each utility throughout the state of California. Figure 16 shows the locations for all the control homes throughout California. Figure 17 is a map of the Los Angeles basin indicating the specific locations of the SCE ADRS and control homes.

Figure 15. ADRS Customer Locations PG&E, SCE, SDG&E

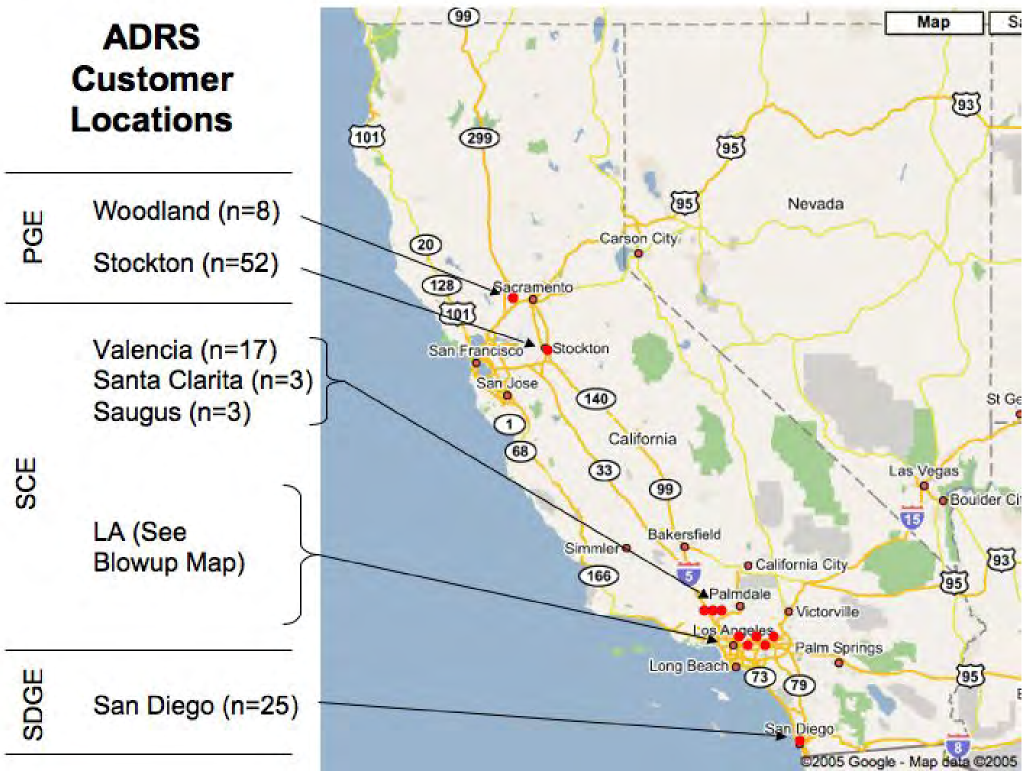


Figure 16. Control Homes Locations PG&E, SCE, SDG&E

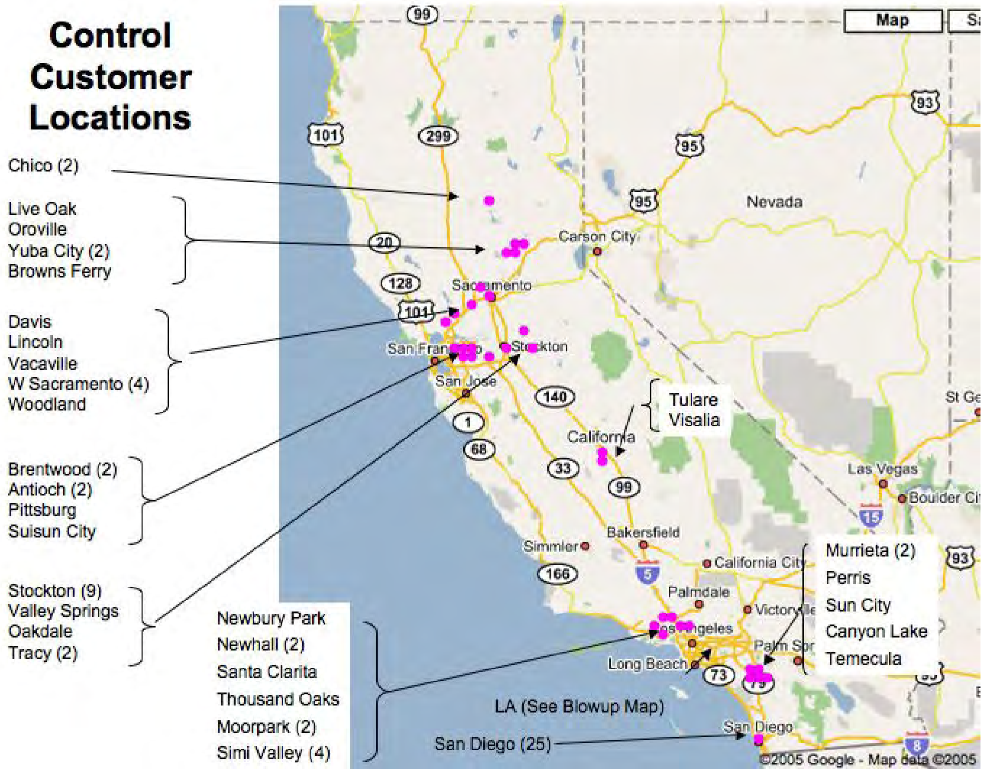
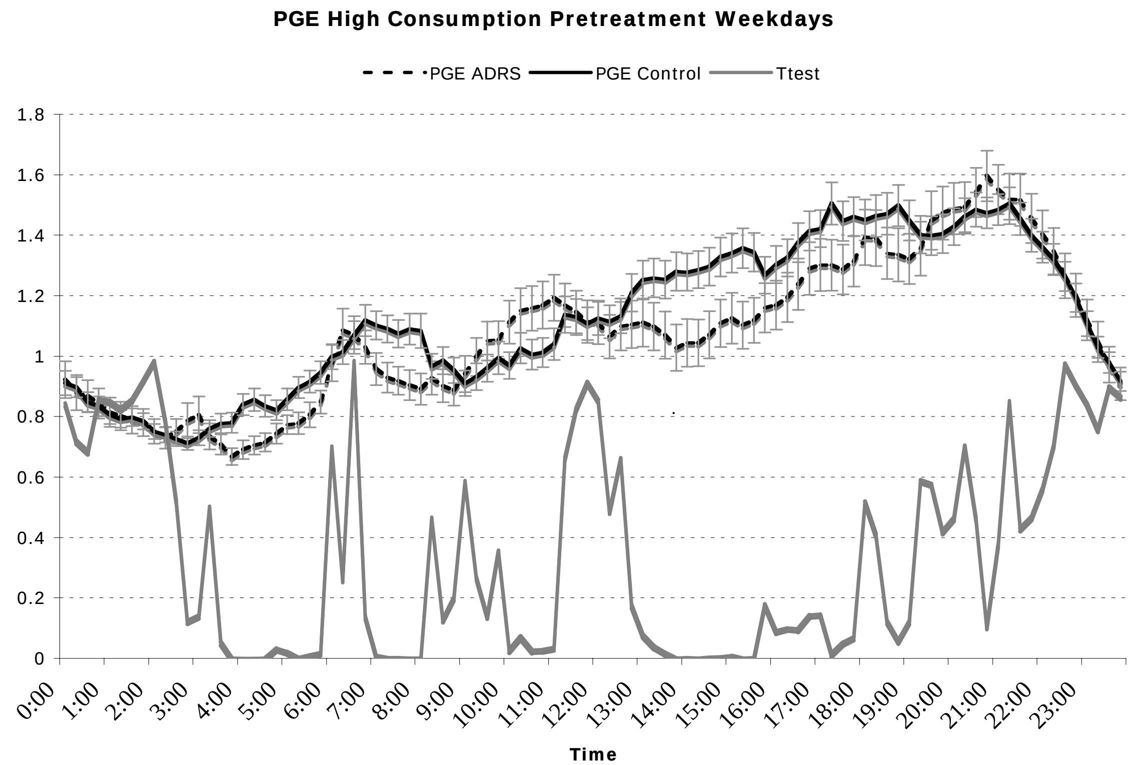


Figure 17. SCE ADRS and Control Homes

PG&E ADRS and control homes are in distinct locations. The PG&E ADRS homes are all located in the Central Valley, while the control homes are dispersed much more broadly: some in the Central Valley, some closer to the foothills of the Sierra Nevada mountains, and some in the East San Francisco Bay area and along the Sacramento River Delta. The Central Valley is generally the hottest region among the regions of interest here, and so it follows that those homes located in the Central Valley would have the highest loads, especially compared to the slightly milder East Bay area. Since ADRS homes were exclusively in the Central Valley, and control homes were spread more widely into milder summer climates, it was hypothesized that the ADRS group's average load would be higher than control's due to higher peak temperatures.

The actual load averages indicated the opposite trend. Figure 18 shows that ADRS customers' average loads in PG&E territory were consistently *lower* than control customers'. RMI cannot explain this behavior with only temperature and interval load data. As such, geography was not considered to be a salient factor for PG&E homes.

Figure 18. PG&E High Consumption Pretreatment Weekdays



A similar analysis was performed on the ADRS and control homes in SCE territory. The SCE ADRS homes were all located in a smaller geographical area than PG&E homes. SCE control homes were confined to the Pomona-Claremont area while SCE ADRS homes were distributed more broadly from East LA to San Bernardino. Even if one were to assume an eastward trend towards higher temperatures, there would be no qualitative inferences one could make in comparing ADRS and control geography. Thus geography was concluded to not have any discernible influence among SCE control and ADRS homes.

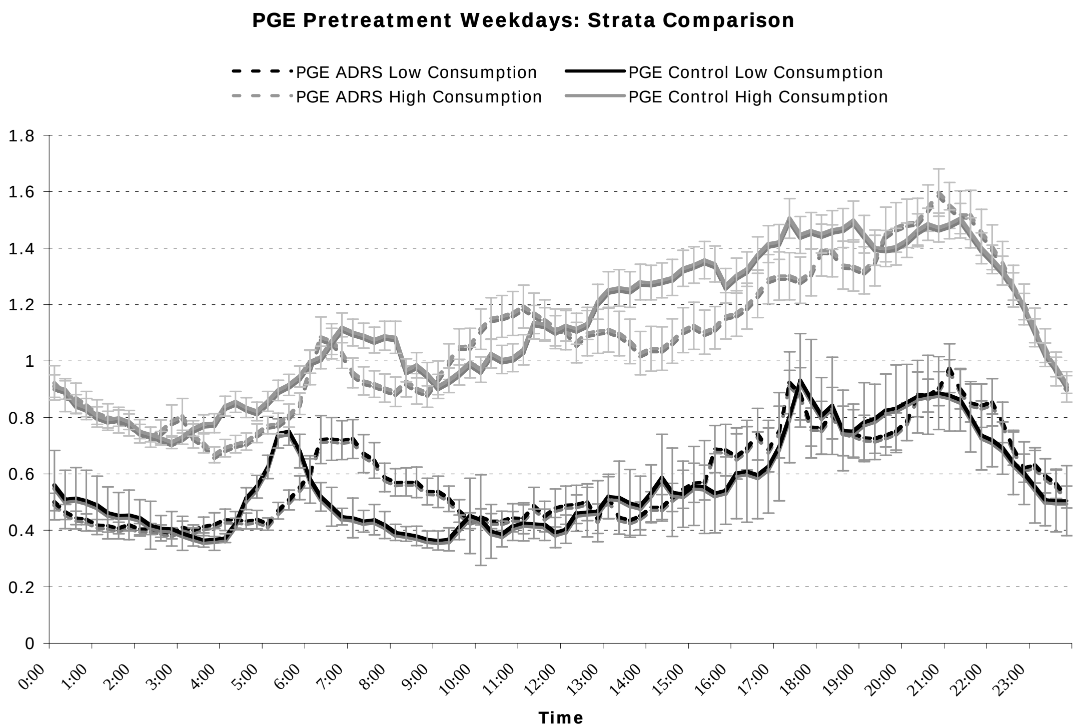
The same conclusion was drawn after examining the geography of SDG&E ADRS and control homes. All SDG&E homes were distributed around the City of San Diego, which was the most confined distribution of any of the utilities. Thus geographic distribution was dismissed as a salient factor in all three utilities.

)<u>Consumption Stratum</u>

The ADRS and control homes had each been divided into strata for high consumption (stratum 2) and low consumption (stratum 1). The third analysis of the pretreatment data was of stratum.

The segmentation of pretreatment data by consumption strata is shown in Figure 19 and Figure 20, for PG&E and SCE, respectively. Both figures indicate consumption differences between control and ADRS customers. For PG&E, there did not appear to be significant differences between ADRS and control in the high consumption with the exception of hours 2 p.m. through 5 p.m., the first three hours of the peak period. During this period, PG&E high consumption ADRS customers consumed less load than control customers. Low consumption ADRS customers consumed about the same as low consumption control homes, except for the morning hours of 6 a.m. to 10 a.m. when ADRS loads were somewhat higher. During the peak period, however, PG&E low consumption ADRS homes consumed about the same load as control homes. Overall, the noise in each stratum was similar to when all data were pooled together (Figure 19) despite the decreased sample size, indicating a relatively stronger signal.

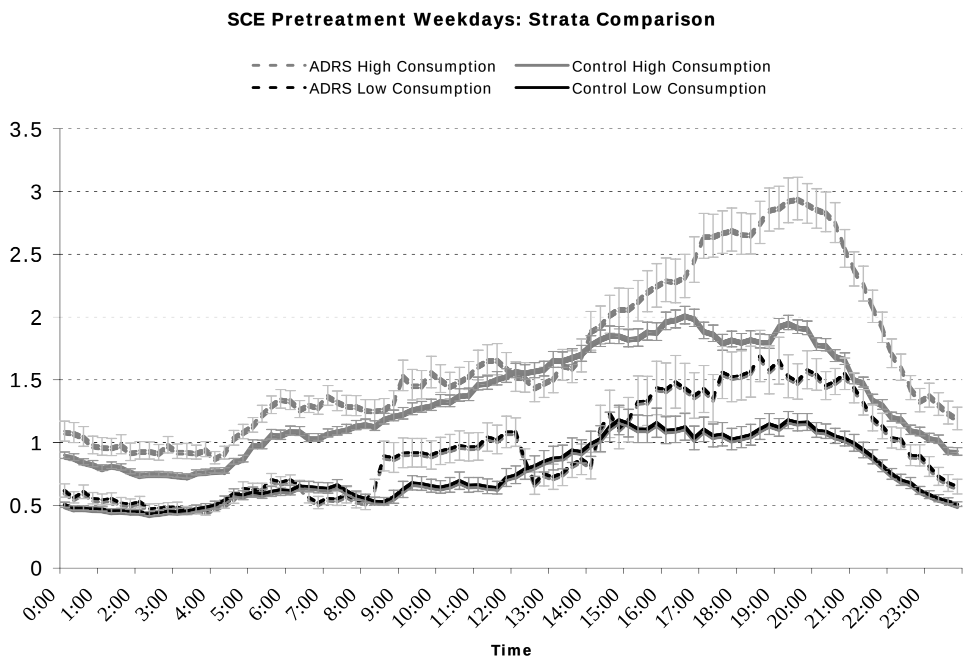
Figure 19. PG&E Strata Comparison



In each ADRS consumption stratum for SCE shown in Figure 20, there was also no significant difference observed for most time intervals except during the last three hours of the peak period, from 4 p.m. to 7 p.m. The differences were significant both for the high and low consumption strata. The high consumption stratum showed a maximum difference of almost 1 kW in the peak period while the low consumption stratum showed a maximum difference of about 0.5 kW between ADRS and control in the peak period. In contrast to the PG&E analysis, ADRS customer loads were greater than control home loads during the last three hours of the peak period for both strata.

The results of pretreatment data division by strata indicated that differences between ADRS and control for both PG&E and SCE were not significant with the exception of the peak periods. For PG&E high consumption customers, ADRS loads were less than control loads during the peak period. For SCE customers, ADRS loads were greater than control loads during most of the peak period. However, the fundamental difficulties inherent to the pretreatment data set, such as small pretreatment sample size and the decline in quantity of pretreatment data throughout May make the conclusion that differences between ADRS and control are significant tenuous. Hence, the data from the summer 2004 weekends were also analyzed.

Figure 20. SCE Strata Comparison



Summer 2004 weekends analysis

ADRS was a technology-based pricing pilot where consumption was reduced according to the way homeowners programmed their ADRS technology, rather than by changing their own behavior. In the ADRS program, weekends featured "off-peak" electricity prices, which were similar to pre-ADRS program electric prices in two respects. First, prices were similar in magnitude to pre-ADRS program prices. Second, electric price behavior during the weekends was similar to pre-ADRS program prices in that they remained constant in magnitude throughout the day. With off-peak electric pricing on weekends, participants would presumably program their ADRS thermostats similar to “default” settings. Hence, summer 2004 weekends load consumption behavior represented the next best proxy for adequate pretreatment data.

Fifteen-minute interval load data for ADRS and control homes were gathered from all three utilities for all weekends and holidays from June-September 2004. The data were then segmented according to high and low consumption strata. Figure 21, Figure 22, and Figure 23 show the high consumption stratum weekend data of the 2004 summer for PG&E, SCE, and SDG&E respectively. ADRS and control home load data on weekend days were gathered for all high consumption homes and average load profiles were calculated from each. Average load profiles for all three utilities exhibit much smoother profiles than the pretreatment data, indicating a more robust data set. All three utilities also exhibit significant differences in consumption between ADRS and control homes during most or all of the peak period. P-values for t-test were close to zero for the majority of time periods in all three utilities. This is consistent with the findings of the pretreatment strata analyses.

Figure 21. PG&E High Consumption Weekends

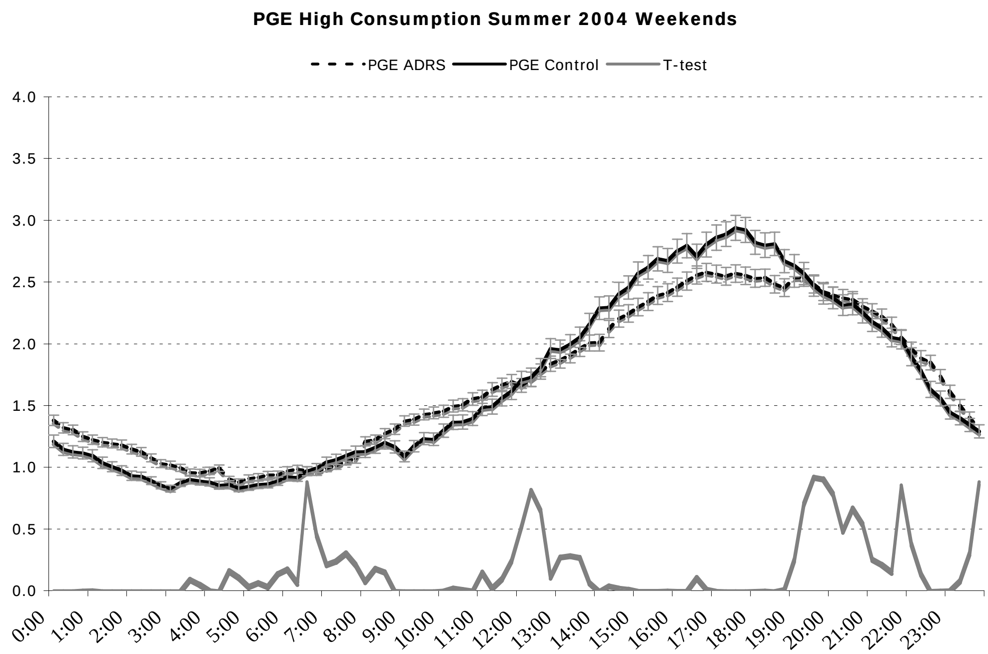


Figure 22. SCE High Consumption Weekends

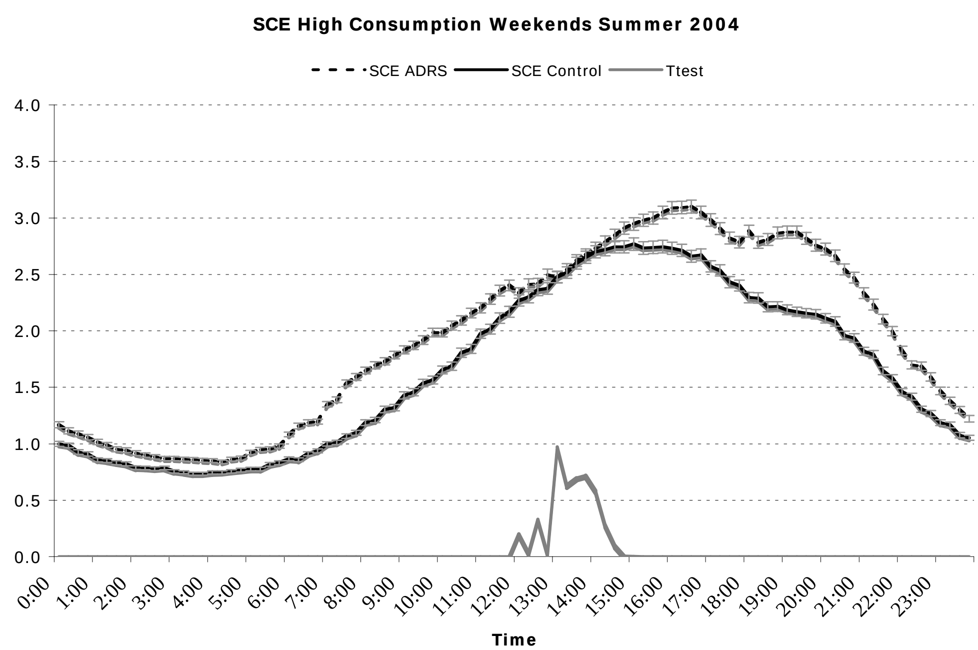
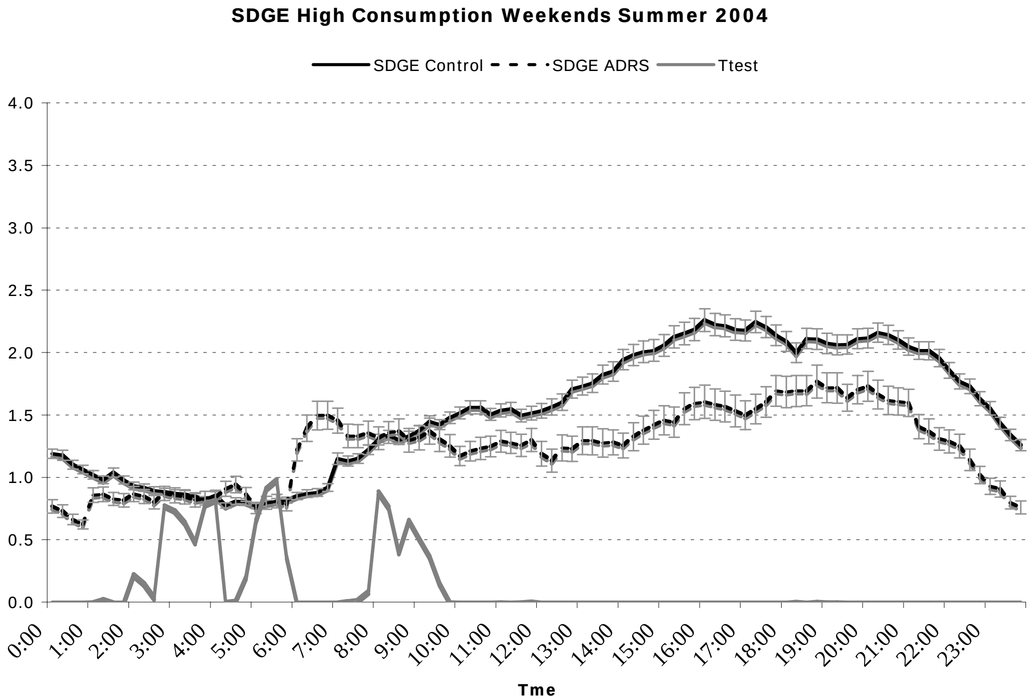


Figure 23. SDG&E Weekends



Given the paucity of homes in the low consumption stratum (see Table 7), the weekends load profiles were even more variable and had more noise. Nevertheless, significant differences were found in similar patterns to the high consumption stratum. PG&E had high t-test p-values values during the morning hours, when ADRS and control had almost identical loads, but those plummeted to zero during the peak period as control homes' average load grew to 1.7 kW beyond the ADRS homes' average load. SCE's low consumption stratum loads featured small differences in the morning that grew to 1 kW during the peak period as the ADRS homes' consumption outpaced the control homes'. SDG&E's low consumption stratum loads were the most jagged, but ADRS and control homes matched each other closely for a majority of the day. During the peak period, control homes' averages exceeded the ADRS homes' load by 0.5 kW.

)<u>Comparison of pretreatment and weekend analyses</u>

The relative magnitude of ADRS consumption compared to control is consistent between the pretreatment weekday and summer 2004 analyses. In the case of SCE, ADRS homes exhibit higher loads compared to control for both high and low strata. This implies that any adjustments for bias in the ADRS participants would be in the positive direction. That is, SCE ADRS load savings relative to control should be increased during the peak period on event and non-event days compared to the simple difference method used in 2004. In the cases of PG&E and SDG&E, ADRS homes exhibit lower load consumption compared to control homes. This implies that adjustments for PG&E and SDG&E ADRS savings using a simple differences comparison would be reduced.

The statistical conclusions between pretreatment and weekends load data, produced similar conclusions. The analysis comparing average weekdays load data prior to GoodWatts technology installation in ADRS homes (pretreatment) showed no statistically significant differences relative to control homes within a specific consumption stratum (high vs. low) with the exception of peak periods. ADRS consumption by strata on Summer 2004 weekends also revealed statistically significant differences relative to the control group, particularly during the peak period.

Given that both pretreatment and weekend data sets were imperfect in their ability to conclusively determine ADRS bias, it would be imprudent to combine the results of both analyses to quantify an “average” difference adjustment for ADRS relative to control homes. From a statistical standpoint, such a combination of results would actually multiply, not reduce, the error of the difference adjustment. A better alternative would be to choose one data set over the other.

Figure 24 and Figure 25 show the average daily differences between ADRS and control homes calculated using pretreatment and summer 2004 weekend data for PG&E and SCE high consumption stratum, respectively. SDG&E was omitted in this comparison given the lack of pretreatment data. The differences adjustment calculated using summer 2004 and pretreatment data match up closely, particularly for PG&E. PG&E’s t-test average p-value calculated over the peak period intervals was 0.56 and 0.43 for the whole day. SCE’s average t-test for SCE was 0.58 for the peak period and 0.33 for the whole day. Thus, we concluded that the differences from pretreatment and summer 2004 weekends were statistically similar. The pretreatment analysis suffered from inadequate data and the summer 2004 weekends data set was generally three times larger than that of the pretreatment. Therefore, the results of the weekend analysis were more statistically robust. The prudent step was to use the summer 2004 weekend data set to quantify the difference adjustment for ADRS by utility and consumption stratum.

2003 monthly billing data

The first source of data for investigating pre-existing consumption differences between ADRS participants and control homes was monthly billing information from 2003. Consumption data from summer 2003 were the most ideal source to use because they were categorically from the “pretreatment” period, before participants were aware of ADRS technology and before participants were subject to critical peak pricing. Unfortunately, detailed fifteen minute interval load data for ADRS and control homes were not available for the summer of 2003. Thus, RMI requested monthly billing data from each of the three utilities.

Because the 2003 billing data and 2004 weekend load data were in different units (kWh per month and kW per 15-minute interval, respectively), Average Daily Usage (ADU) was used as a common metric for a consistent basis of comparisons. Figure 26, Figure 28, and Figure 30 show the ADU calculated for May through September from the 2003 monthly billing data for PG&E, SCE, and SDG&E respectively. The monthly kWh consumed from each home was divided by the number of days in the billing cycle to produce an ADU for the month for each home. PG&E’s and SDG&E’s 2003 ADU show control homes’ average ADU consistently above the ADRS homes’ average ADU but with no statistically significant differences. SCE’s 2003 ADU shows the ADRS homes’ ADU consistently above the control homes’ average ADU with no statistically significant differences between the two.

Figure 24. PG&E difference comparison using pretreatment and weekends data

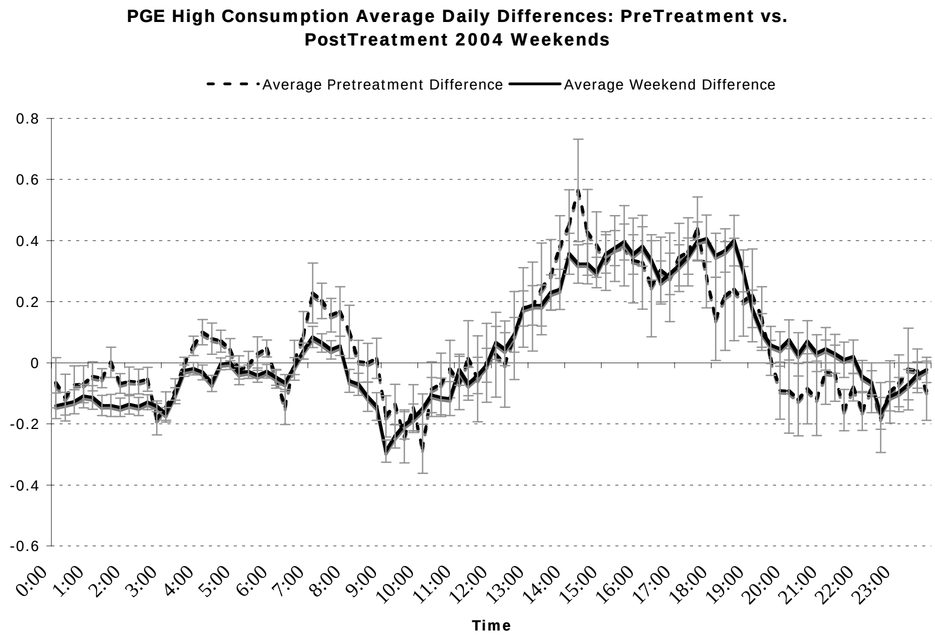
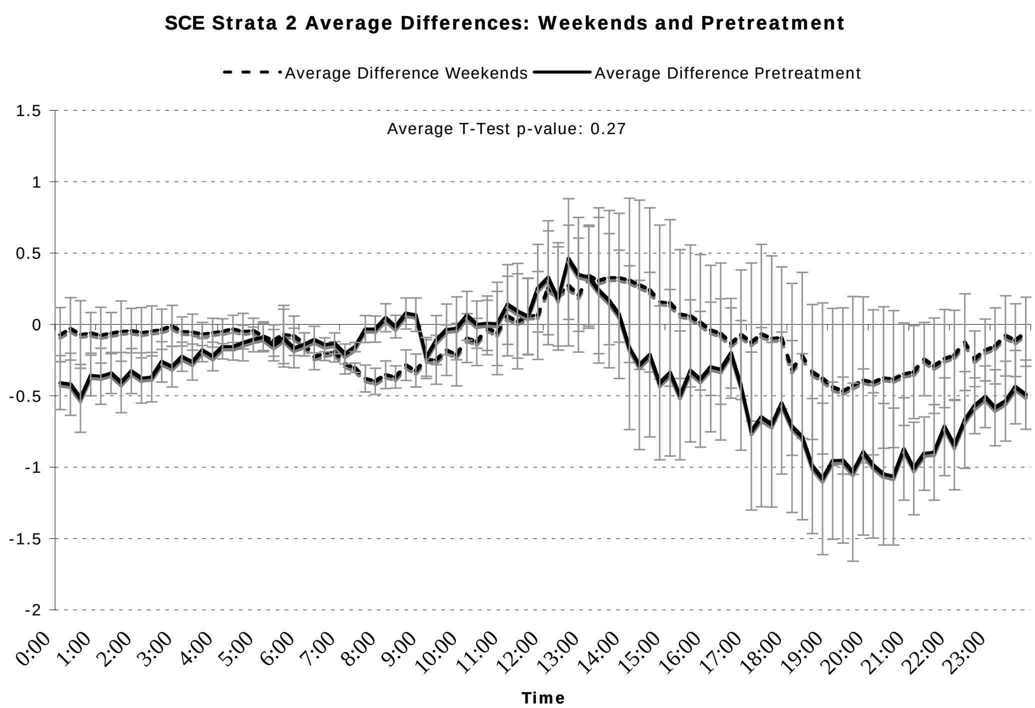


Figure 25. SCE difference comparison using pretreatment and weekends data



The 2003 ADU data represent a true snapshot of pretreatment, but of much lower resolution. The hypothesis was that if the 2004 weekend data were truly different from pretreatment, this difference might be shown in a comparison of 2003 monthly ADU and 2004 monthly ADU based on the weekend data. For example, if the 2004 PG&E control homes' ADU were significantly *lower* than ADRS homes' ADU while 2003 PG&E control homes' ADU were higher (but insignificant), this would support the hypothesis that 2004 weekend data were different from pretreatment data, and therefore might *not* be a suitable proxy for pretreatment. Note that it is only appropriate to compare relative patterns in the data between years, since there were many factors (such as temperature), which varied between summers.

Figure 27, Figure 29, and Figure 31 show the monthly ADU calculated from the summer 2004 weekends data. There were several steps involved in converting the weekend load data to ADU values for each month. First, all the weekend load data were divided by month, and then the total daily consumption was calculated for each home on each day of weekend data. Then, each home's daily consumption values were averaged in each month, resulting in a monthly ADU for each home. All the monthly ADUs for each home were averaged together by month, resulting in the ADUs represented in the figures.

Figure 26. PG&E High Consumption 2003 ADU

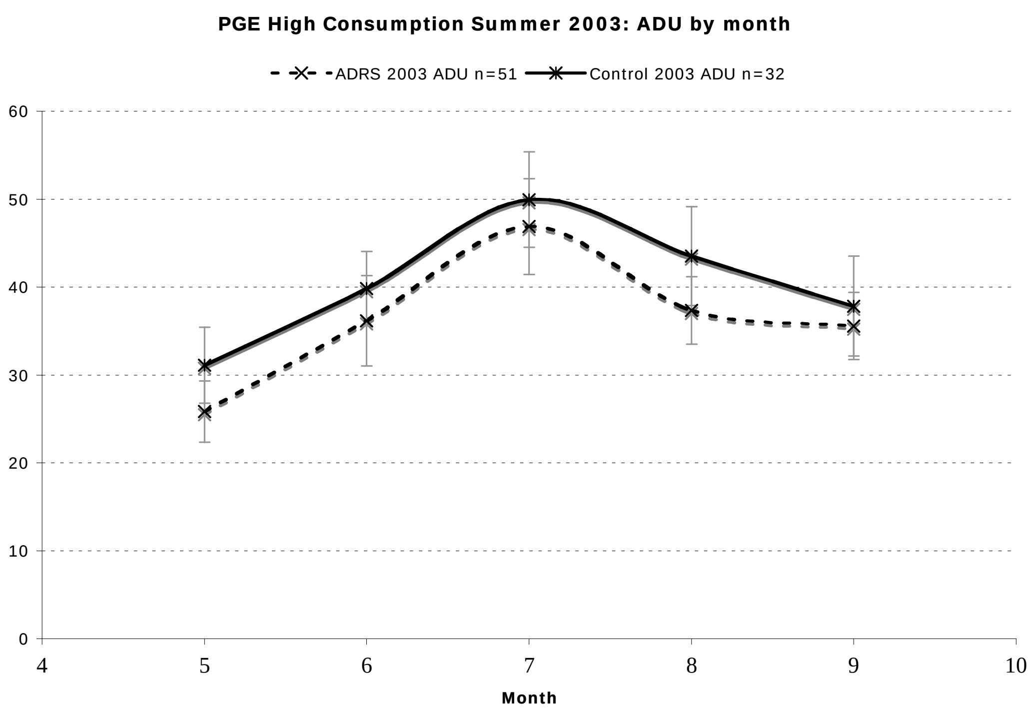


Figure 27. PG&E High Consumption 2004 Weekend ADU

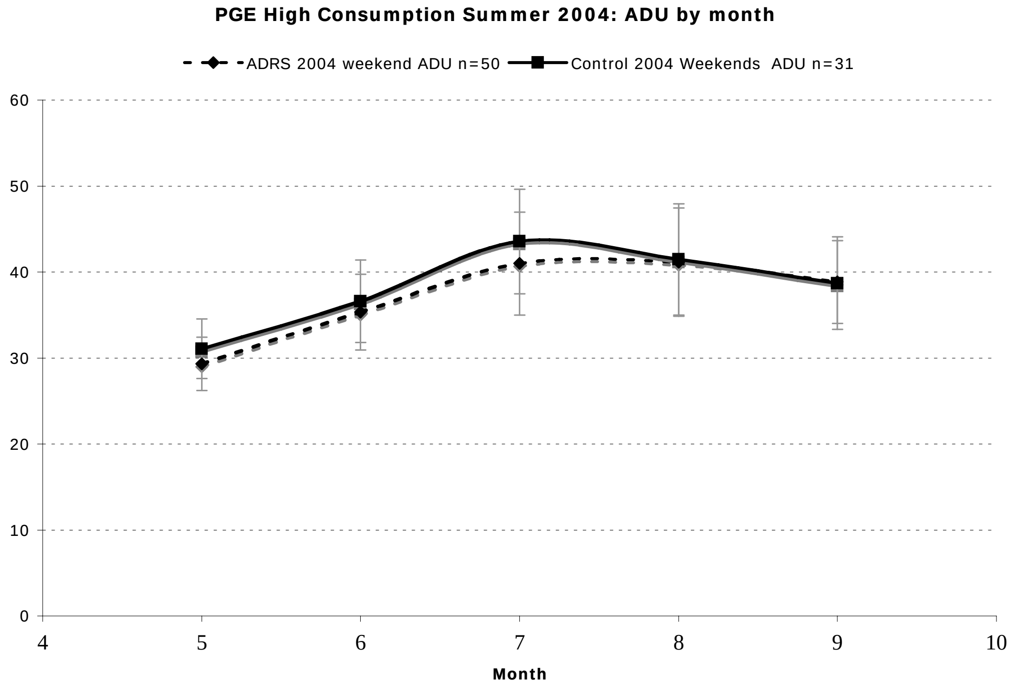


Figure 28. SCE High Consumption 2003 ADU

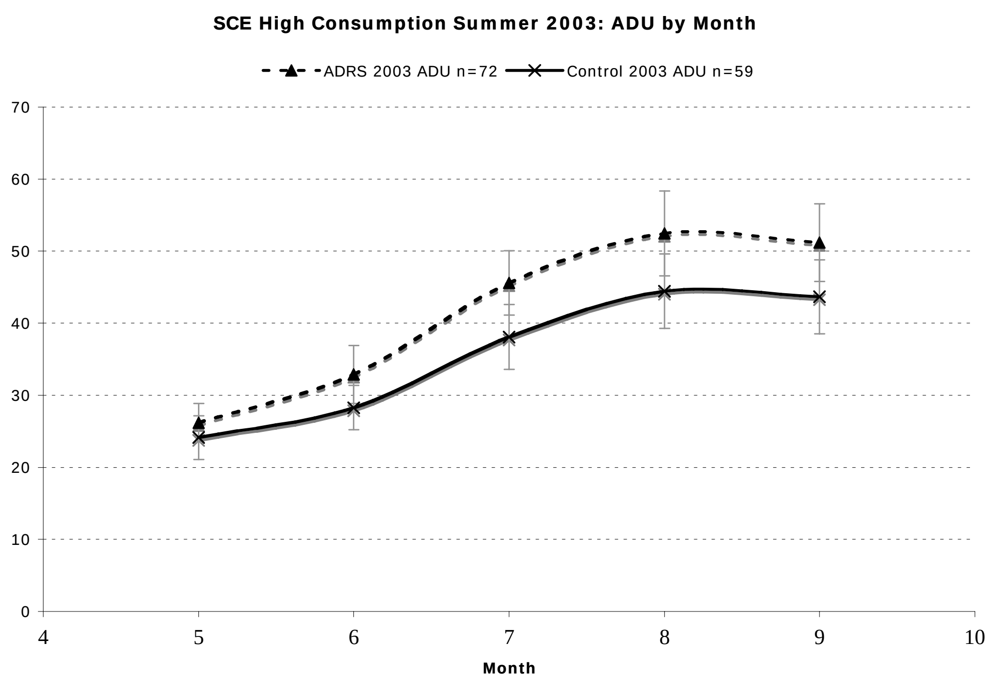


Figure 29. SCE High Consumption 2004 ADU Weekends

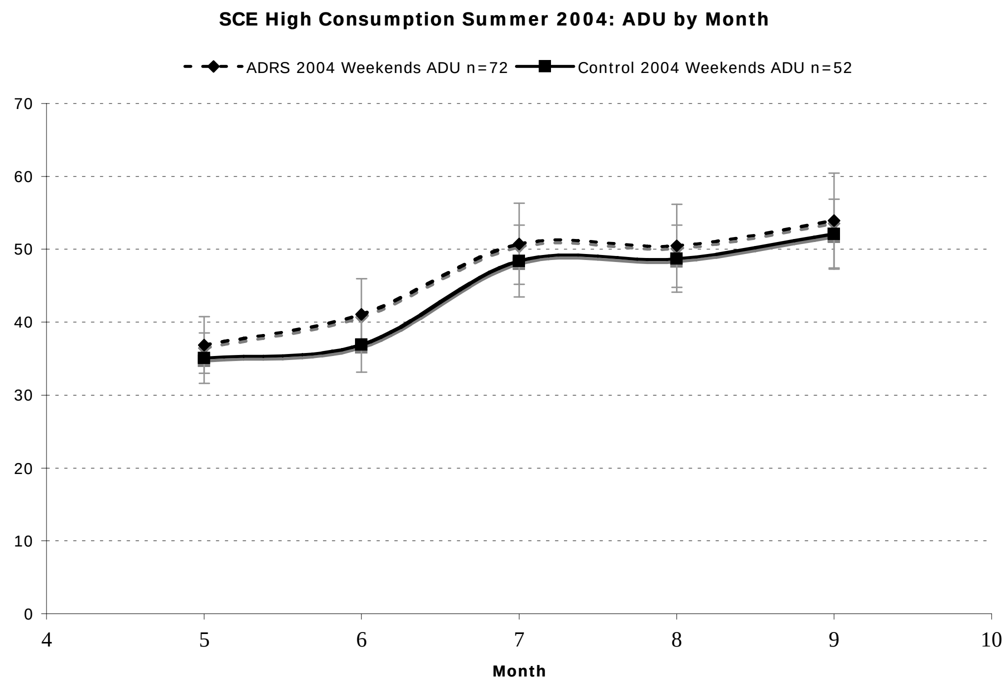


Figure 30. SDG&E 2003 ADU

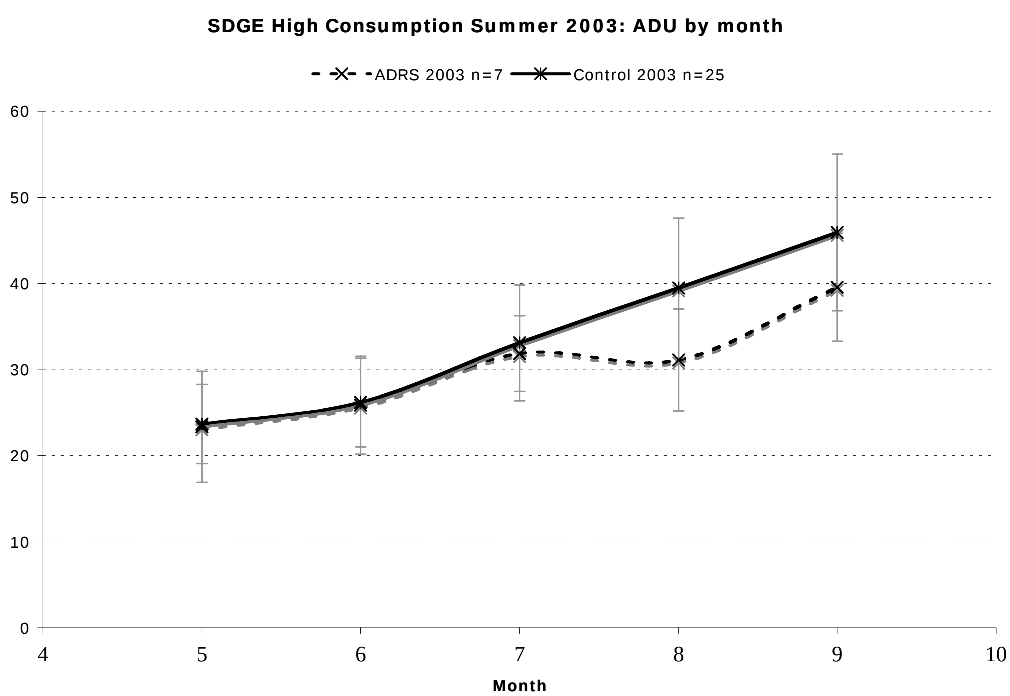
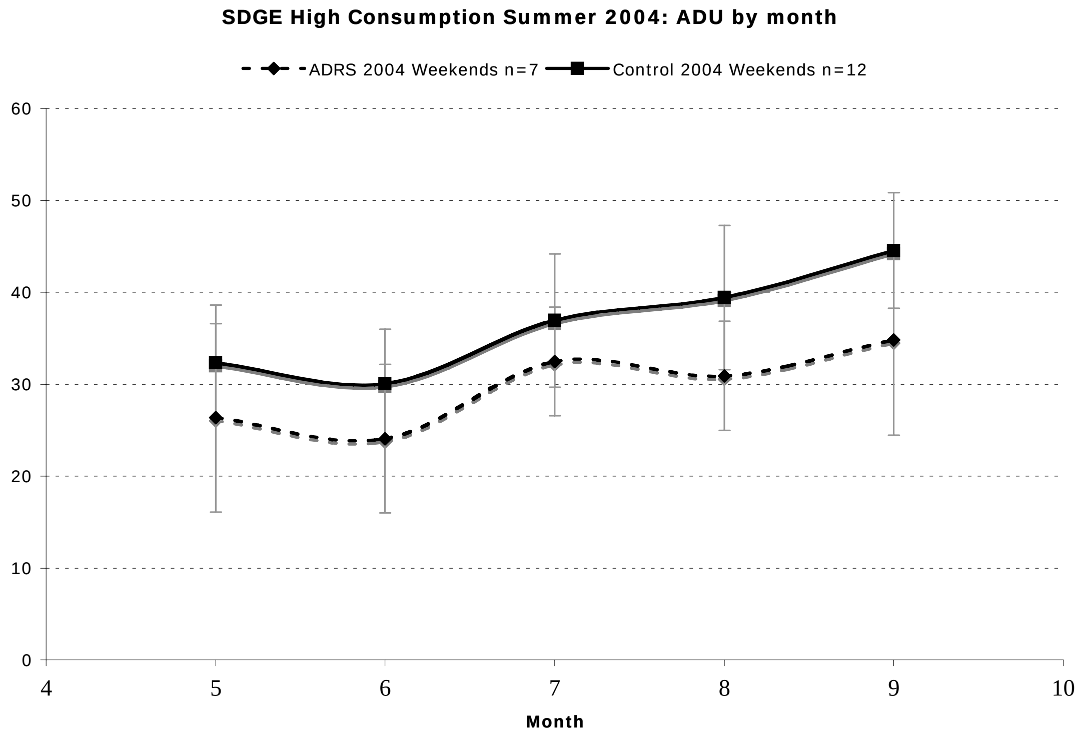


Figure 31. SDG&E 2004 ADU



The figures produce consistent results with pretreatment and weekend analyses. In PG&E and SDG&E territories, the average ADU for control homes is slightly more than the ADU for ADRS homes, though not significantly different. SCE's 2004 ADUs for control homes are consistently less than ADRS's average ADUs, but not significantly different. All these observations conform exactly with the observations from the 2003 ADU data.

Given these similarities in comparison, 2004 weekend data could not be considered different from 2003 monthly billing data based on this analysis. The summer of 2004 weekend data had passed the two tests used to validate them. Thus, the summer 2004 weekend data were accepted as a proxy for pretreatment data and the differences were calculated for each of the three utilities between control and ADRS by stratum.

Applying the differences adjustment

Figure 32, Figure 33, and Figure 34 show the result of applying the differences adjustment for high consumption stratum of PG&E, SCE, and SDG&E, respectively, using the difference values calculated from summer 2004 weekends data. The charts show the ADRS load profiles for each utility before and after application of the differences adjustment on event and non-event days. The thickest black line close to the x-axis plots the differences adjustment in each chart for each utility.

PG&E difference adjustments are small (± 0.1 kW) through most of the day, and increase to 0.4 kW during the peak period. SCE differences are all negative, reflecting the higher consumption of ADRS customers compared to control customers. SCE’s peak period differences start at 0.2 kW and decrease steadily through the period down to -0.2 kW. SDG&E's differences vary between -0.5 kW at dawn to 0.6 kW during the peak period. SDG&E's differences are not quite as smooth as PG&E's or SCE's, which is likely an effect of the small number of high consumption stratum ADRS homes (see Table 7).

Figure 32. PG&E High Consumption Adjustment Chart

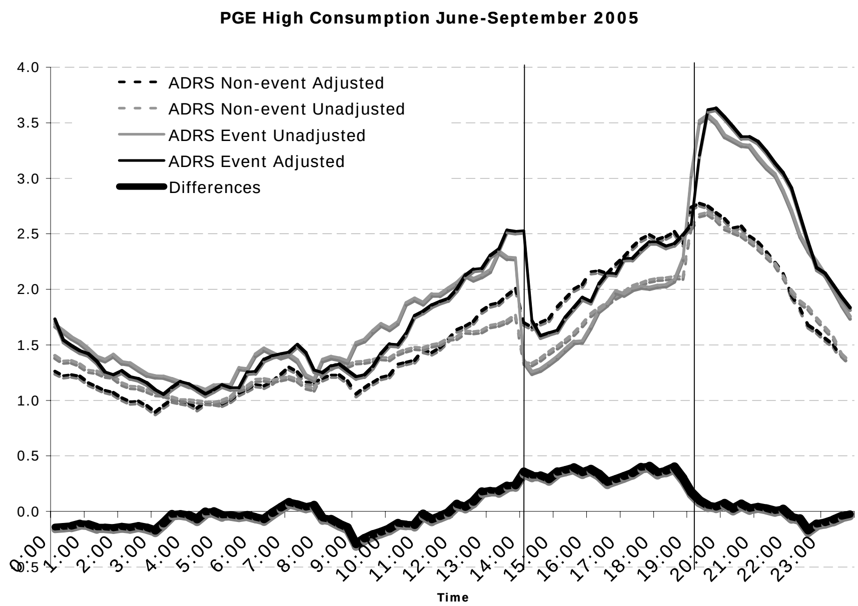


Figure 33. SCE High Consumption Adjustment

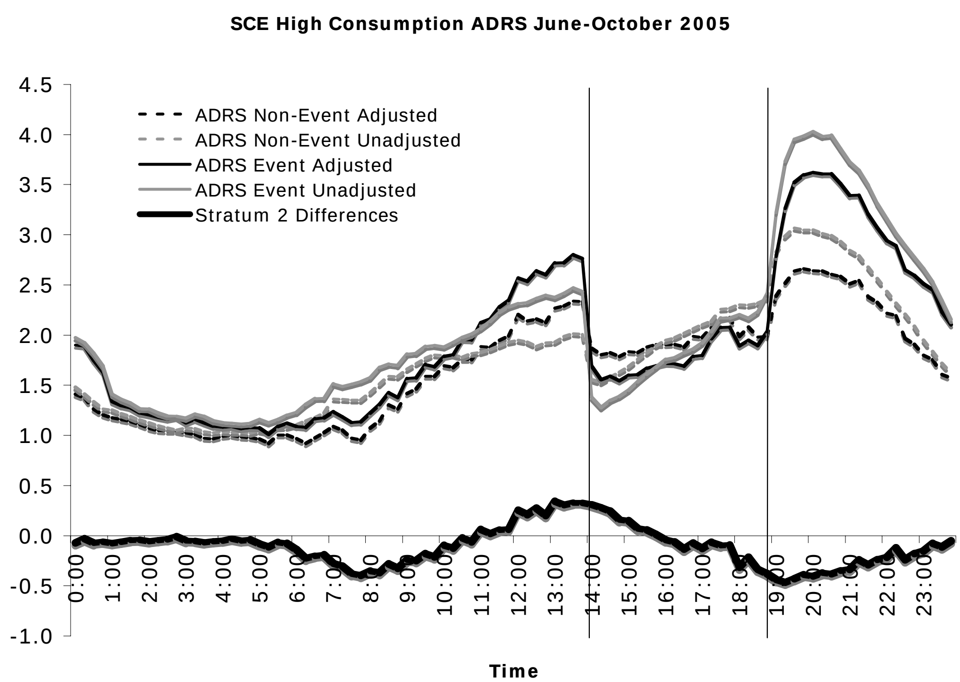


Figure 34. SDG&E High Consumption Adjustment

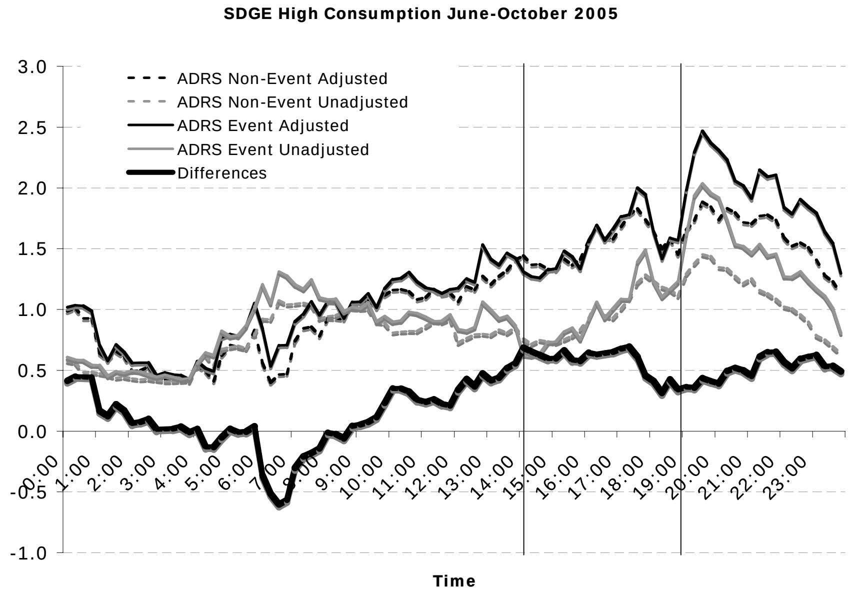


Figure 35. PG&E Low Consumption Adjustment

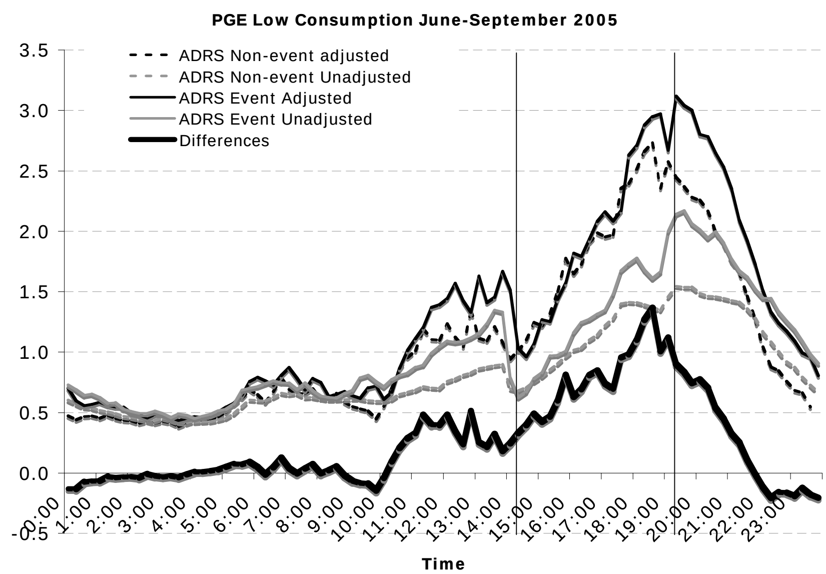


Figure 35 is a chart of PG&E's low consumption stratum adjustment, which is typical of the low consumption stratum adjustments for SCE and SDG&E as well. The differences curve is noisier and more extreme than any of the high consumption stratum adjustments. Differences are nearly zero throughout the early morning while the peak period rises from 0.3 kW up to 1.4 kW. The low consumption stratum differences exhibit these extreme characteristics as a result of a limited data set (see Table 7). In PG&E service territory, there are only two control homes in the low consumption stratum.

ADRS selection bias analysis summary

In summary, RMI turned to four sources of information in our investigation into ADRS customer selection bias: ADRS program recruiting and welcome materials, interval meter data during ADRS "pretreatment" weekdays prior to ADRS (GoodWatts) technology installation, interval meter data during 2004 summer weekends, and monthly kWh billing meter data from summer 2003. The first source was qualitative and the last three sources were quantitative evaluations. The three quantitative sources provided us with data in varying levels of detail and quality.

The three quantitative data sources all produced consistent patterns in the orientation of ADRS participant consumption compared to control homes. Moreover, the orientations were utility specific. For PG&E and SDG&E, the pretreatment data, summer 2004 weekends, and 2003 monthly billing data all concluded that ADRS customers consumed less load than control customers. For SCE, the quantitative analyses all concluded that ADRS customers consumed more load than control customers.

The fifteen minute interval data available for pretreatment and summer 2004 weekends analyses revealed furthermore that differences in control and ADRS customers consumption were statistically significant during the hours of 2 p.m. to 7 p.m., when ADRS customers are charged higher electric rates during event and non-event days, according to the CPP-F experimental rate schedule. Because the pretreatment data were particularly problematic and the summer 2004 weekends data were more robust, and because the utility specific differences between the pretreatment and weekends data were similar, RMI decided to apply the adjustments resulting from the weekends analysis to all subsequent ADRS pilot load impact evaluations.

Table 10. Summary of hourly peak period selection bias adjustments, based on control-ADRS customers' average consumption differences using summer 2004 weekends data

	2 p.m.	**3 p.m.**	**4 p.m.**	**5 p.m.**	**6 p.m.**	**Average**
PGE Difference High Stratum	0.325	0.370	0.318	0.368	0.355	0.347
SCE Difference High Stratum	0.247	0.074	-0.078	-0.095	-0.311	-0.033
SDGE Difference High Stratum	0.640	0.606	0.641	0.609	0.372	0.574
PGE Difference Low Stratum	0.366	0.575	0.745	0.842	1.186	0.743
SCE Difference Low Stratum	-0.736	-0.842	-0.959	-0.958	-0.824	-0.864
SDGE Difference Low Stratum	-0.358	-0.309	-0.203	-0.236	-0.250	-0.271
Statewide High Only	0.248	0.163	0.071	0.067	-0.077	0.094
Statewide ALL	0.260	0.202	0.116	0.070	-0.050	0.119

Although the average load differences between control and ADRS customers were statistically significant, the magnitude of the differences was small with few exceptions (Table 10). The PG&E service territory high consumption differences varied little from 0.3 kW across the 5 hour peak period. Control-ADRS differences for PG&E low consumption customers varied more between, 0.4

kW at 2 p.m. increasing gradually to 1.2 kW at 6 p.m. For SCE customers, high consumption differences ranged from a high of 0.25 kW at 2 p.m. to a low of -0.3 kW at 6 p.m. Control-ADRS differences for SCE low consumption customers varied between -0.7 kW to -0.9 kW across the peak period. Finally, for SDG&E high consumption customers, control-ADRS differences were consistently 0.6 kW from 2 p.m. to 5 p.m., dropping to 0.3 kW at 6 p.m. Differences for SDG&E low consumption customers also exhibited small variation, ranging from 0.4 kW at 2 p.m. to 0.2 kW at 6 p.m.

The statewide differences adjustments were derived from the weighted average of the difference adjustments calculated for each utility. Because the orientation of ADRS load consumption compared to control were in opposite directions between SCE and PG&E and SDG&E, the resulting statewide difference adjustments were close to zero (Table 10).

Appendix D
Low Consumption ADRS
2005 and 2004 Load Impact Results Charts

Figure 36: 2005 Statewide low consumption homes: average of event days load curves

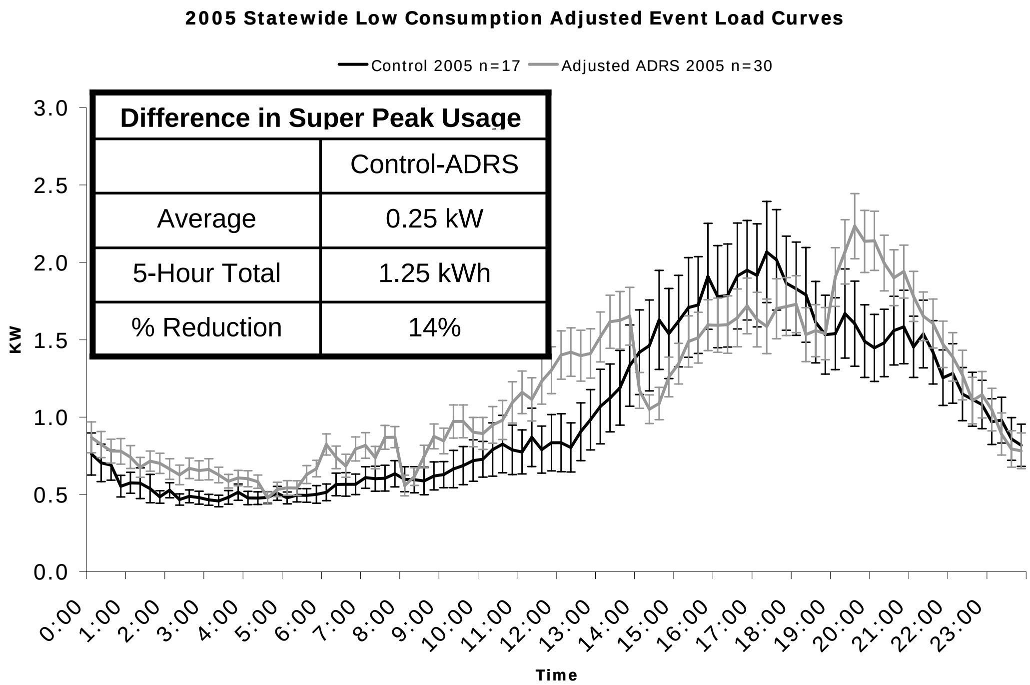


Figure 37: 2005 Statewide low consumption homes: average of non-event weekdays load curves

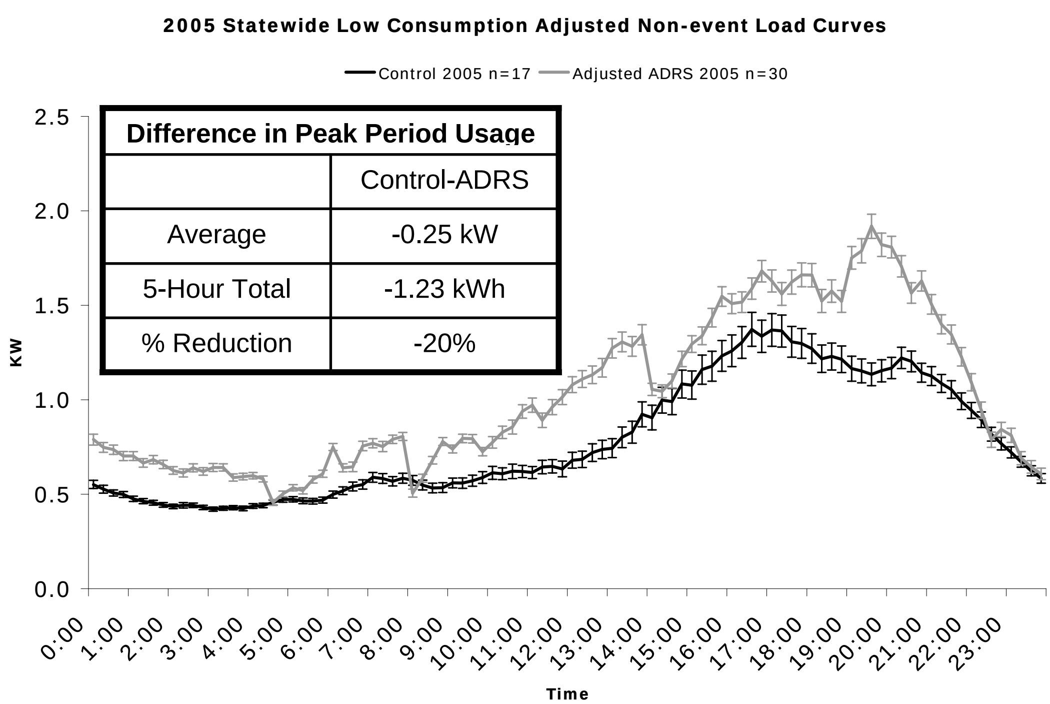


Figure 38: 2005 Statewide low consumption homes: hourly Super Peak period load reductions

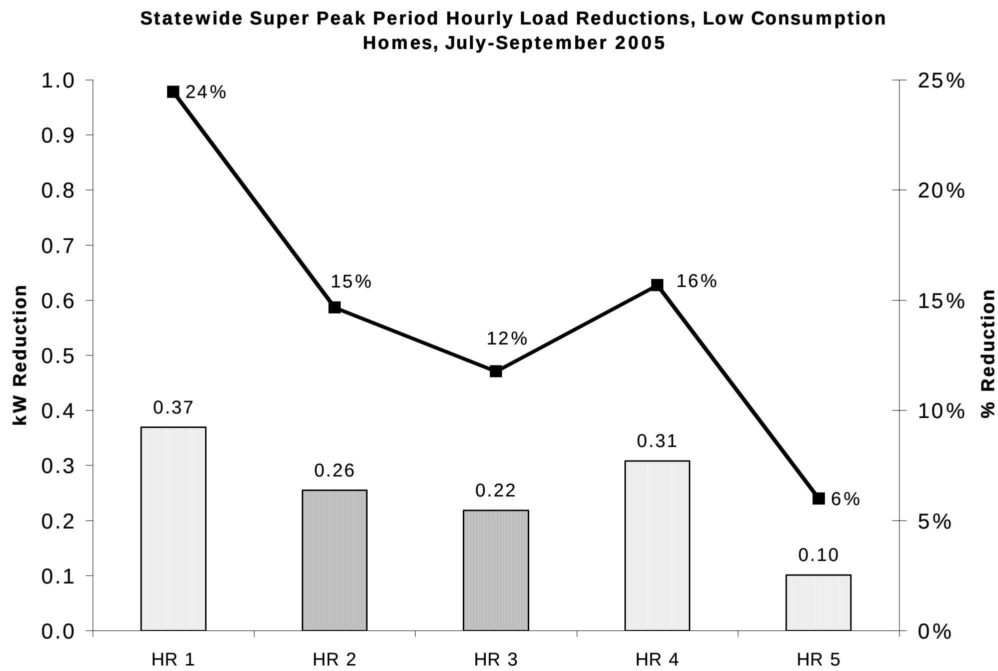


Figure 39: 2005 Statewide low consumption homes: hourly peak period load reductions

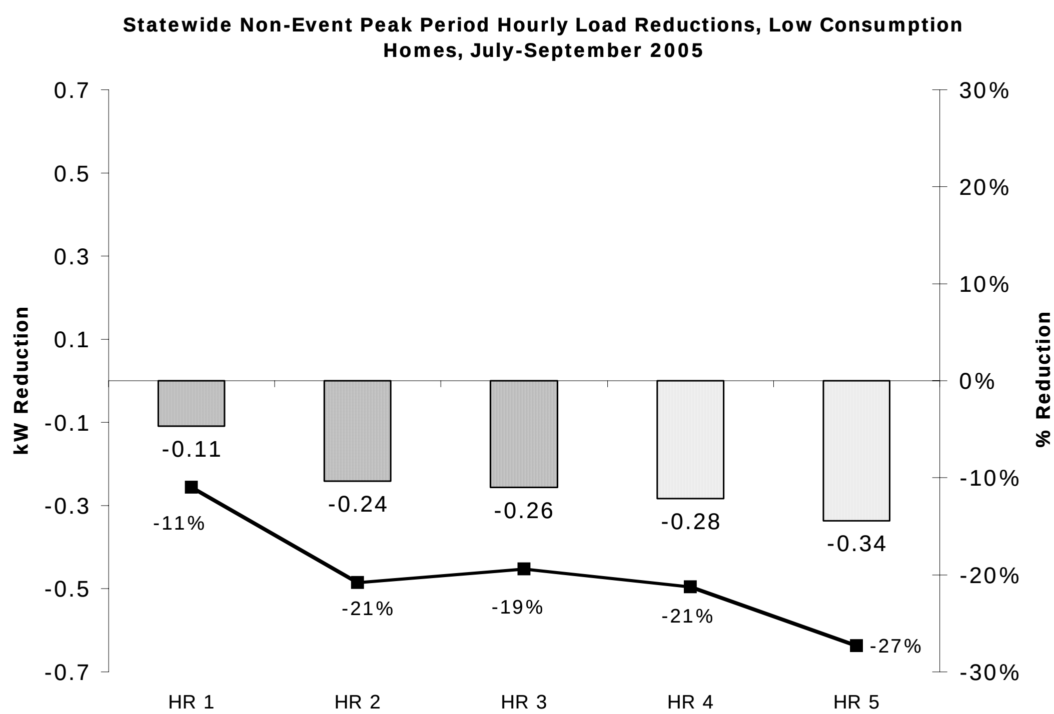


Figure 40: PG&E low consumption homes: average of event days load curves, 2005

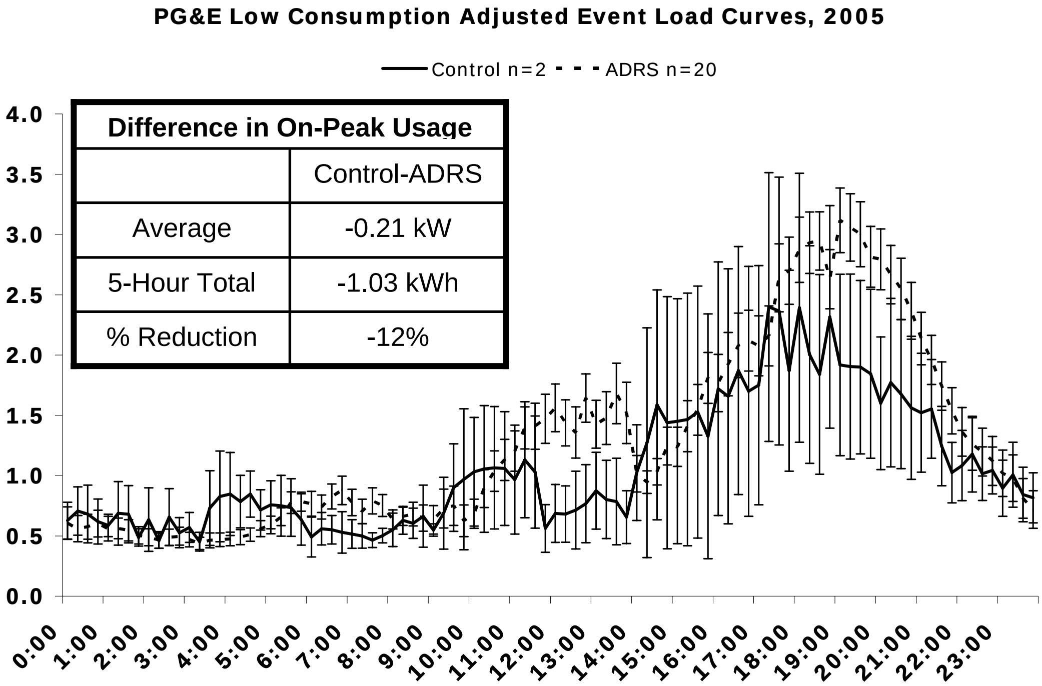


Figure 41: PG&E low consumption homes: average of non-event weekdays load curves, 2005

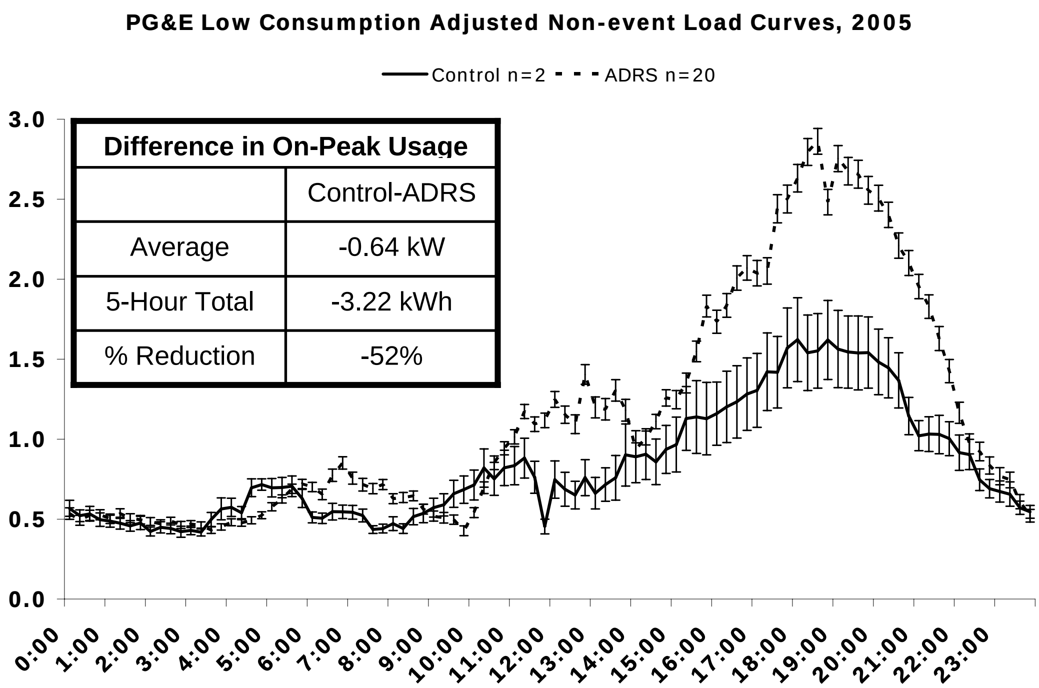


Figure 42: PG&E low consumption homes: hourly Super Peak period load reductions, 2005

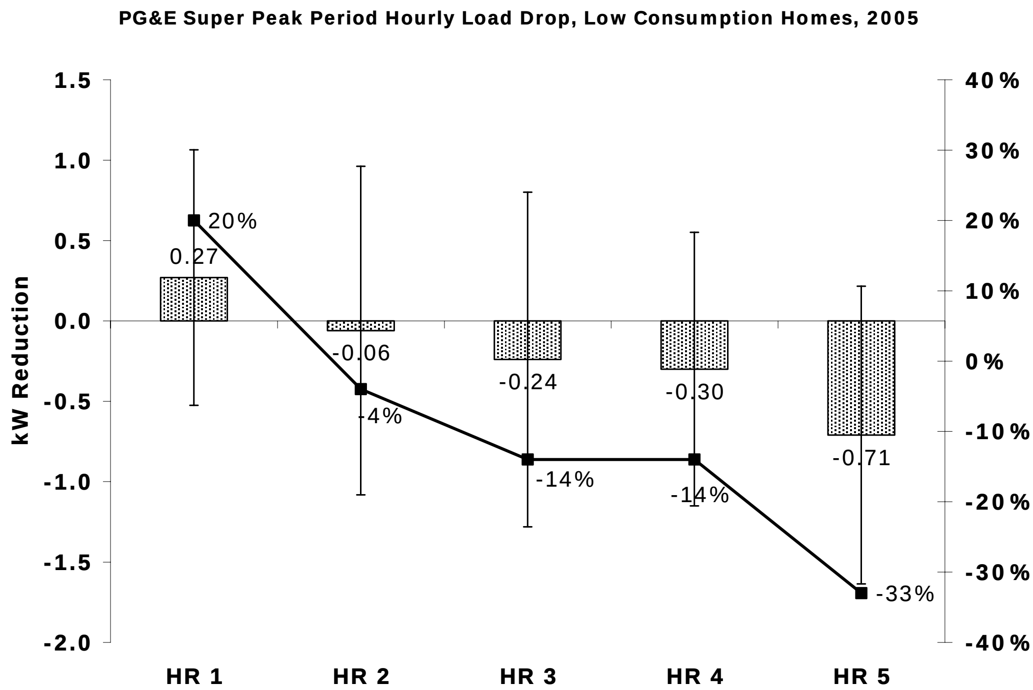


Figure 43: PG&E low consumption homes: hourly peak period load reductions, 2005

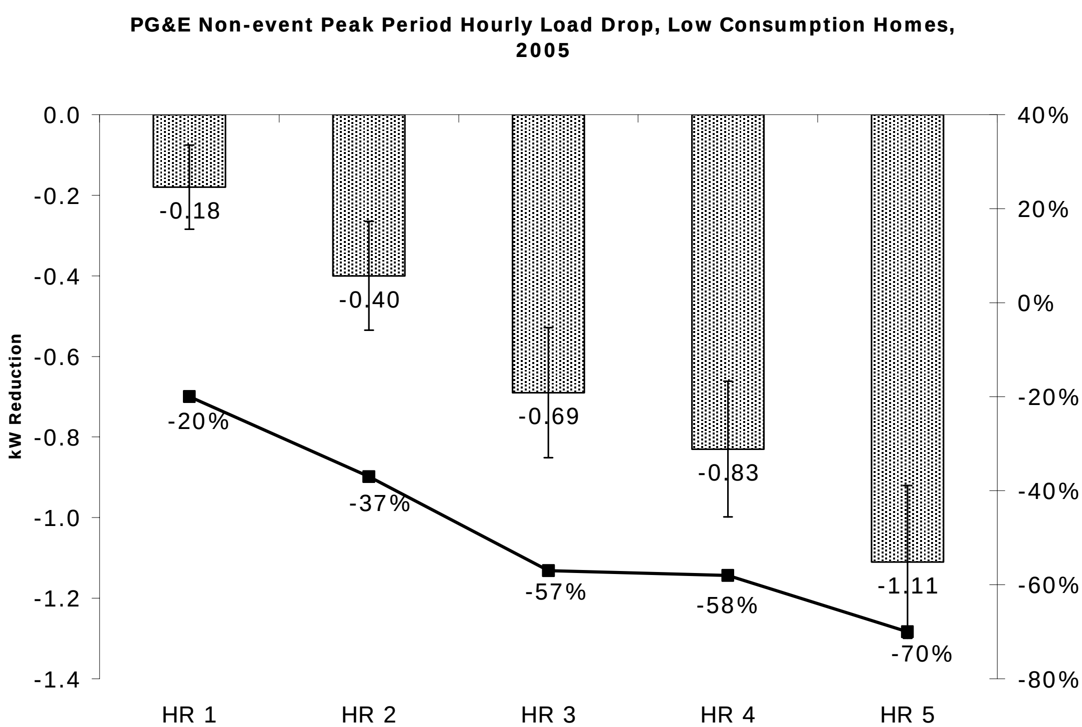


Figure 44: SCE low consumption homes: average of event days load curves, 2005

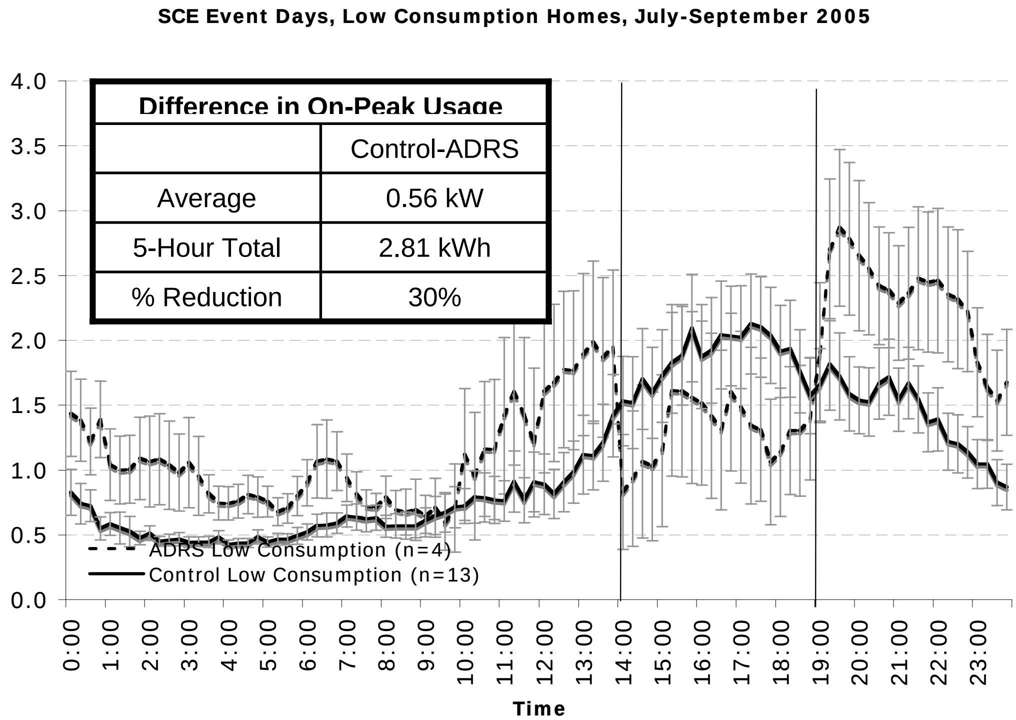


Figure 45: SCE low consumption homes: average of non-event weekdays load curves, 2005

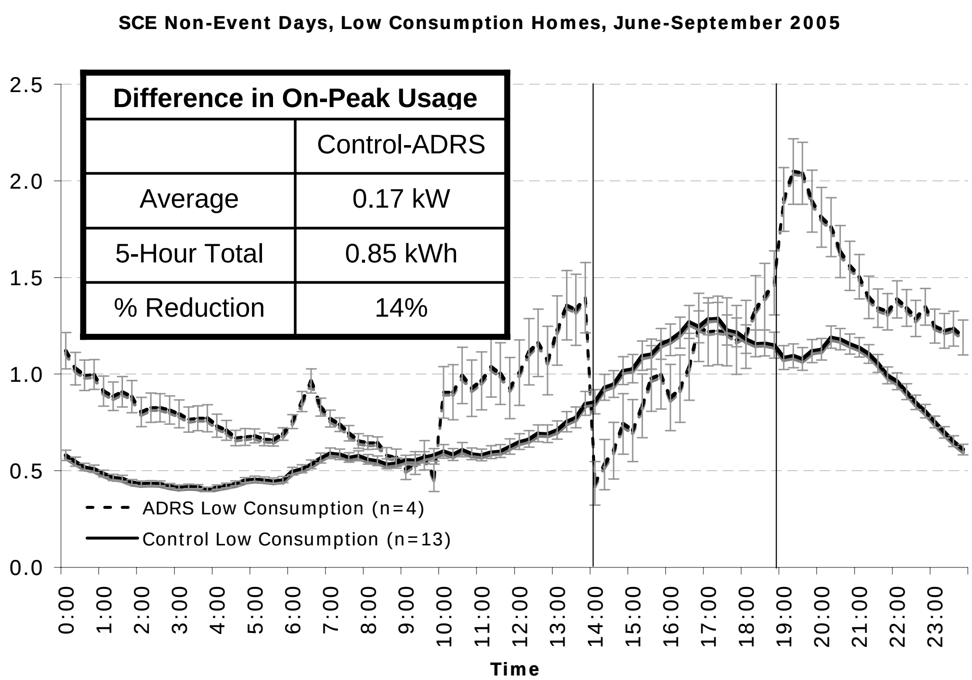


Figure 46: SCE low consumption homes: hourly Super Peak period load reductions, 2005

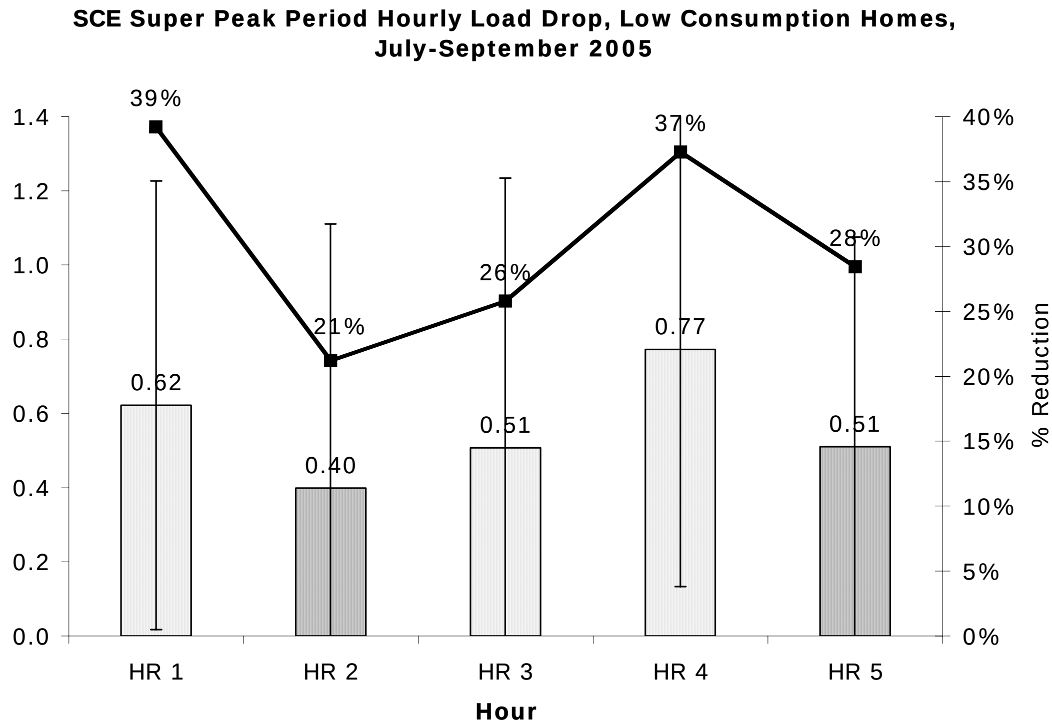


Figure 47: SCE low consumption homes: hourly peak period load reductions, 2005

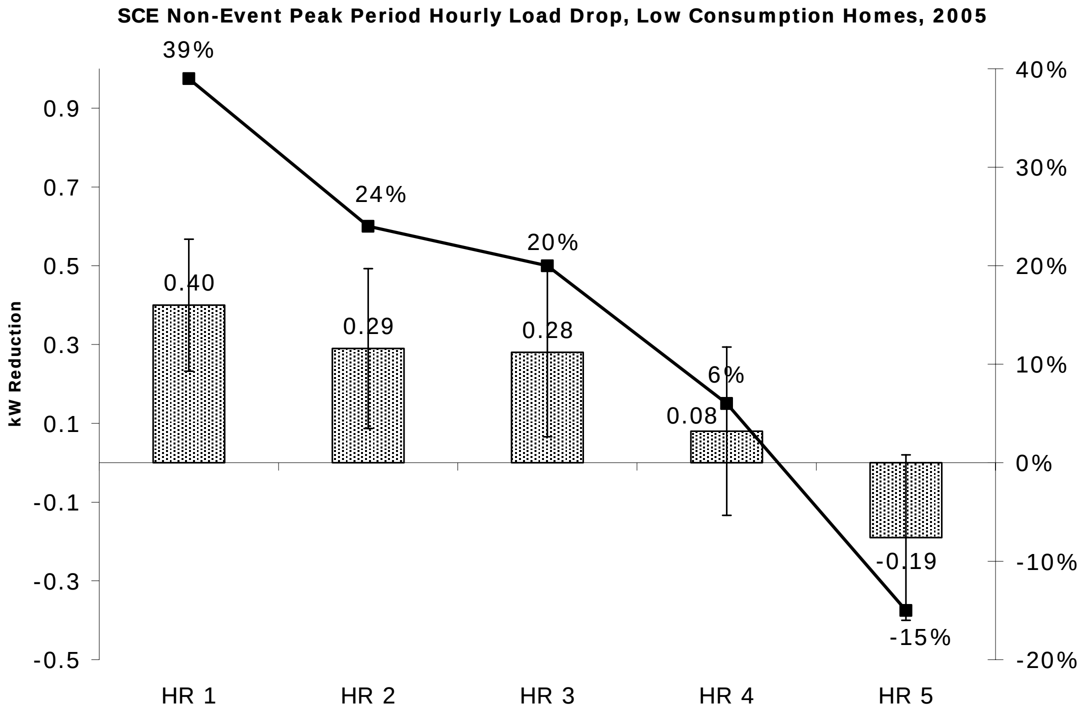


Figure 48: SDG&E low consumption average of event days load curves, 2005

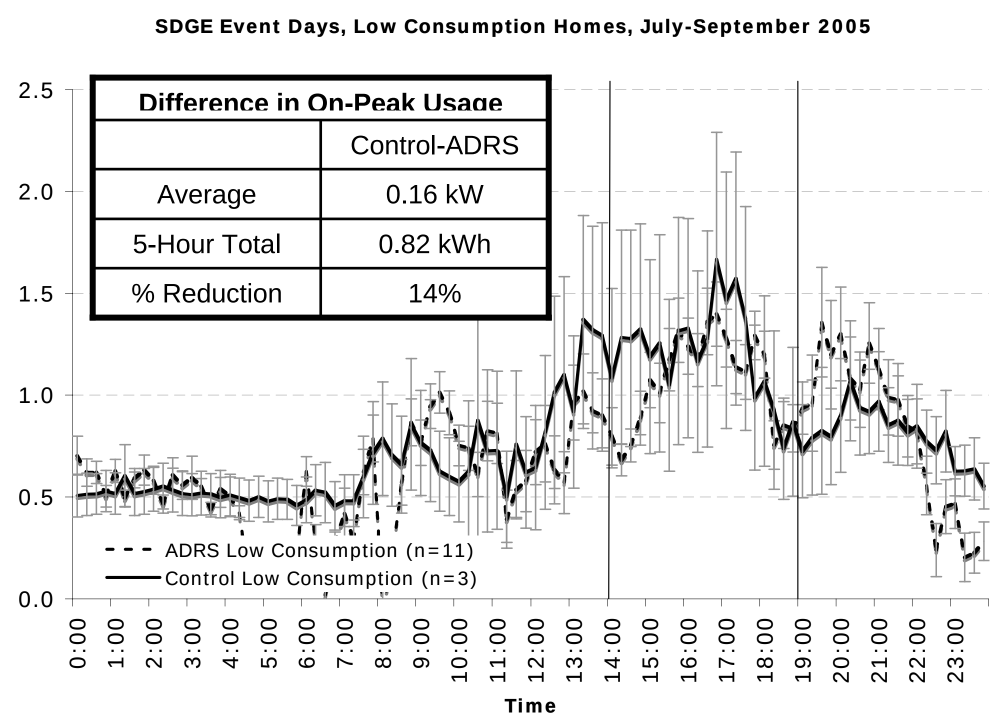
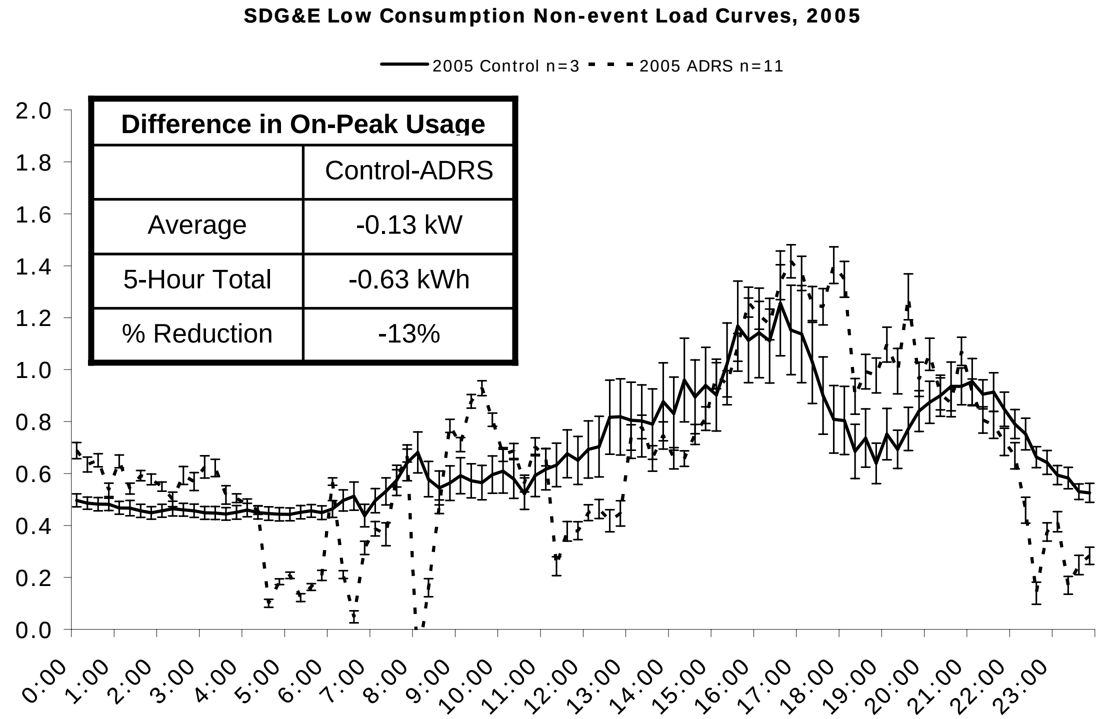


Figure 49: SDG&E low consumption average of non-event weekdays load curves, 2005



Difference in On-Peak Usage	
	Control-ADRS
Average	-0.13 kW
5-Hour Total	-0.63 kWh
% Reduction	-13%

Figure 50: SDG&E low consumption homes: hourly Super Peak period load reductions, 2005

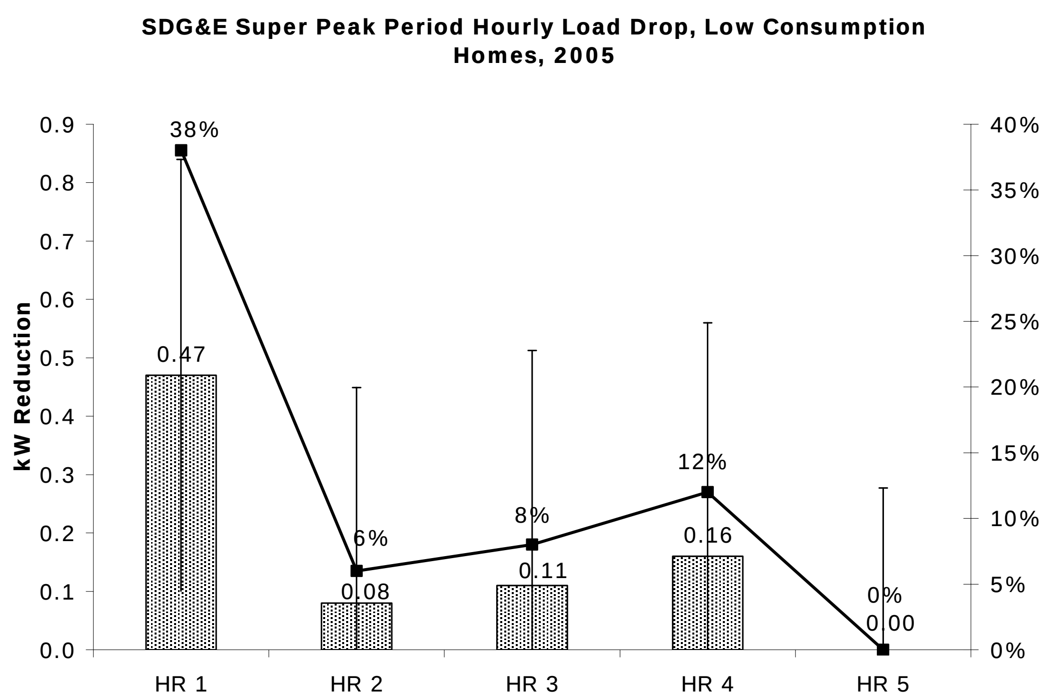


Figure 51: SDG&E low consumption homes: hourly peak period load reductions, 2005

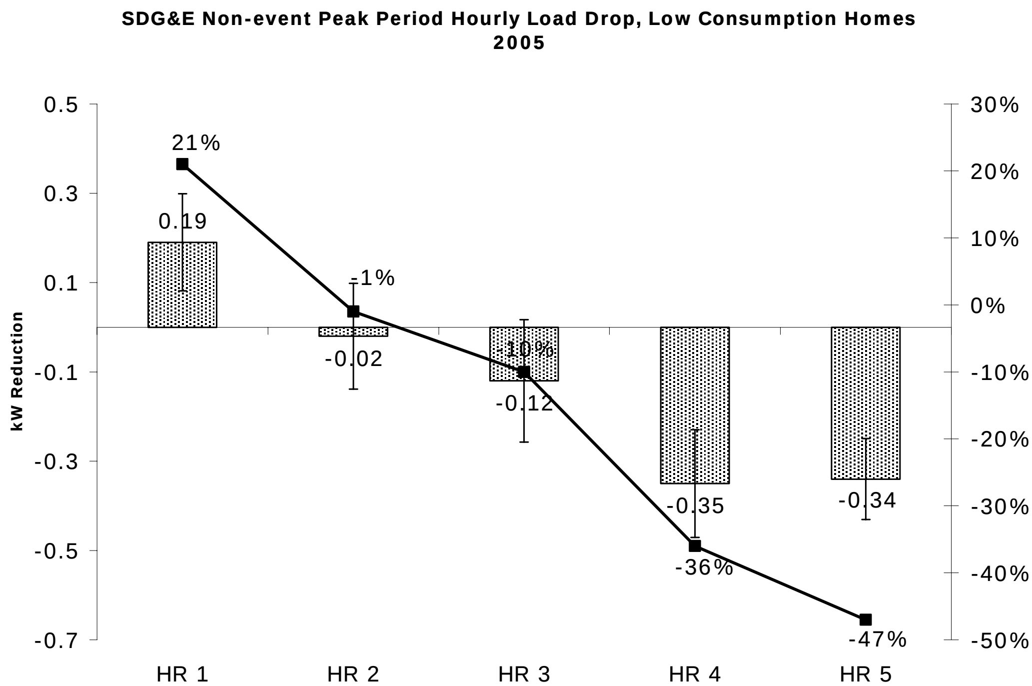


Figure 52: Comparison of 2005 and 2004 statewide low consumption homes: average of event days load curves

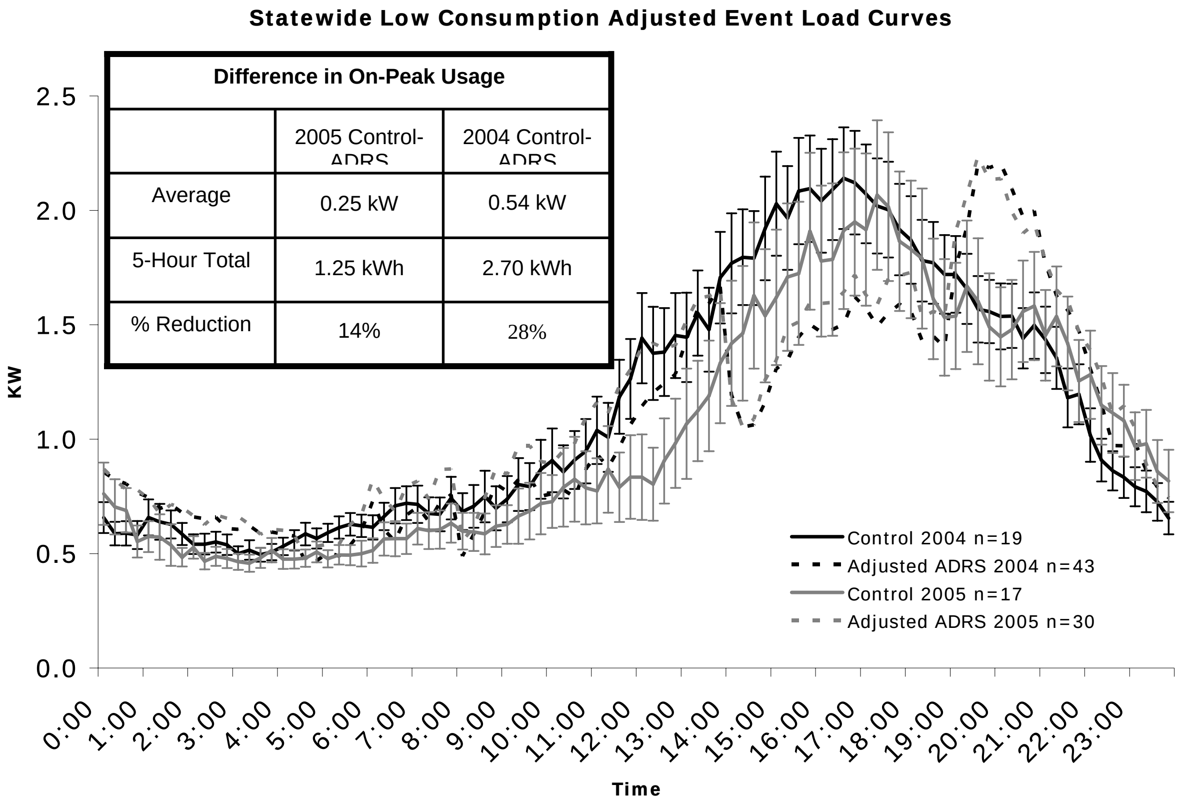


Figure 53: Comparison of 2005 and 2004 statewide low consumption homes: average of non-event weekdays load curves

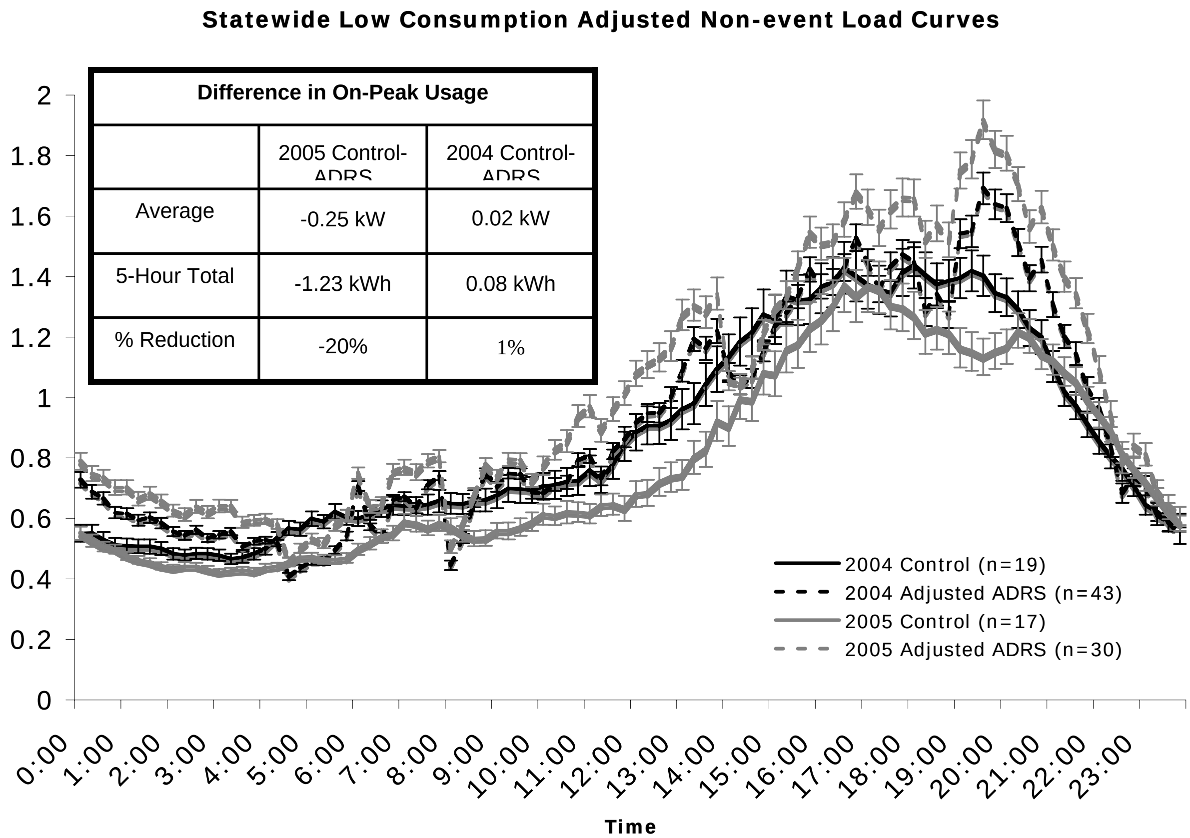

Statewide Low Consumption Adjusted Non-event Load Curves
Difference in On-Peak Usage
2005 Control-ADRS
2004 Control-ADRS
Average
-0.25 kW
0.02 kW
5-Hour Total
-1.23 kWh
0.08 kWh
% Reduction
-20%
1%
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0:00
1:00
2:00
3:00
4:00
5:00
6:00
7:00
8:00
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
22:00
23:00
Time
2004 Control (n=19)
2004 Adjusted ADRS (n=43)
2005 Control (n=17)
2005 Adjusted ADRS (n=30)

Figure 54: Comparison of 2005 and 2004 PG&E low consumption homes: average of event days load curves

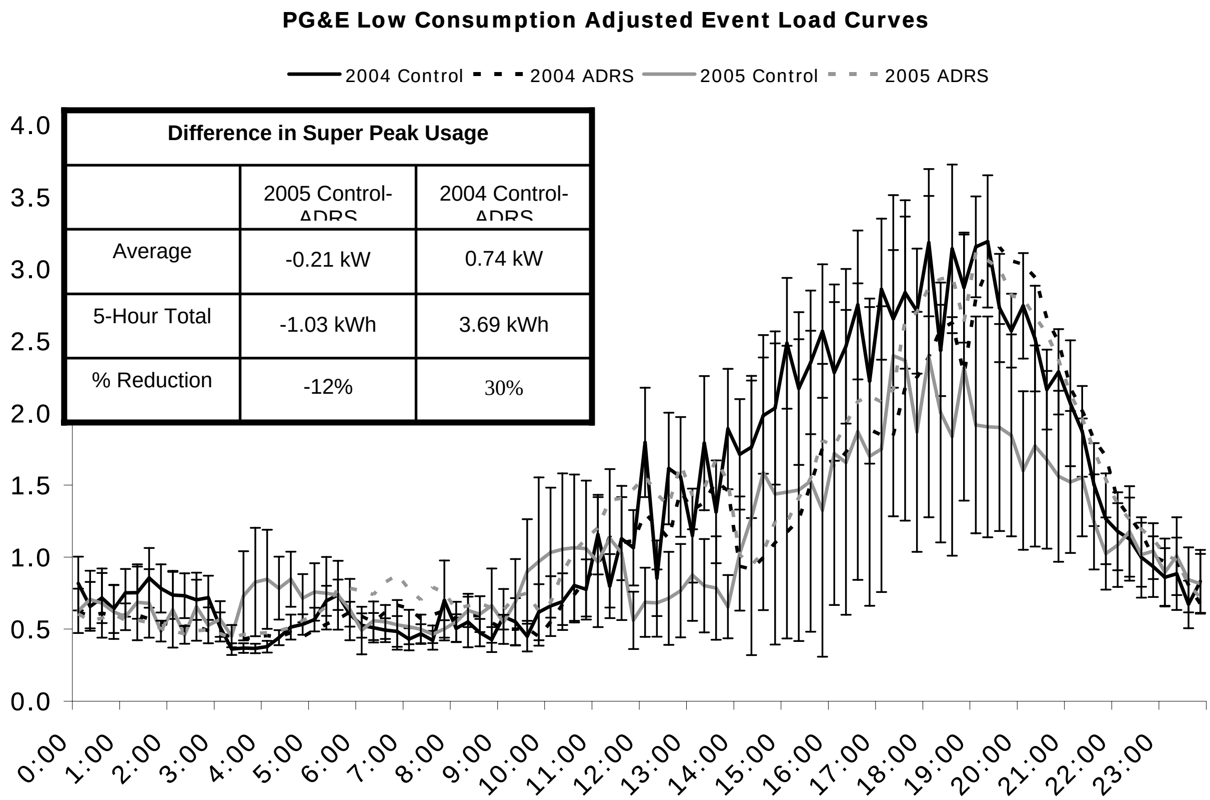


Figure 55: Comparison of 2005 and 2004 PG&E low consumption homes: average of non-event weekdays load curves

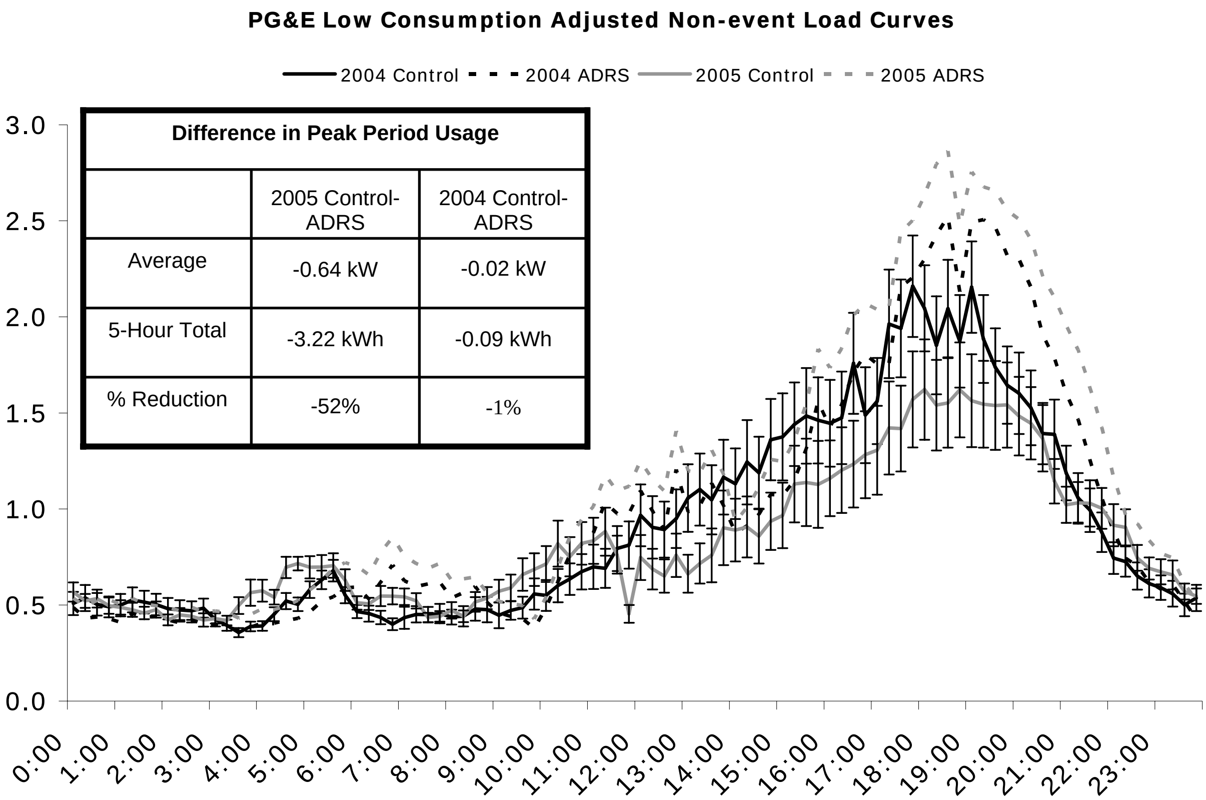


Figure 56: Comparison of 2005 and 2004 SCE low consumption homes: average of event day load curves

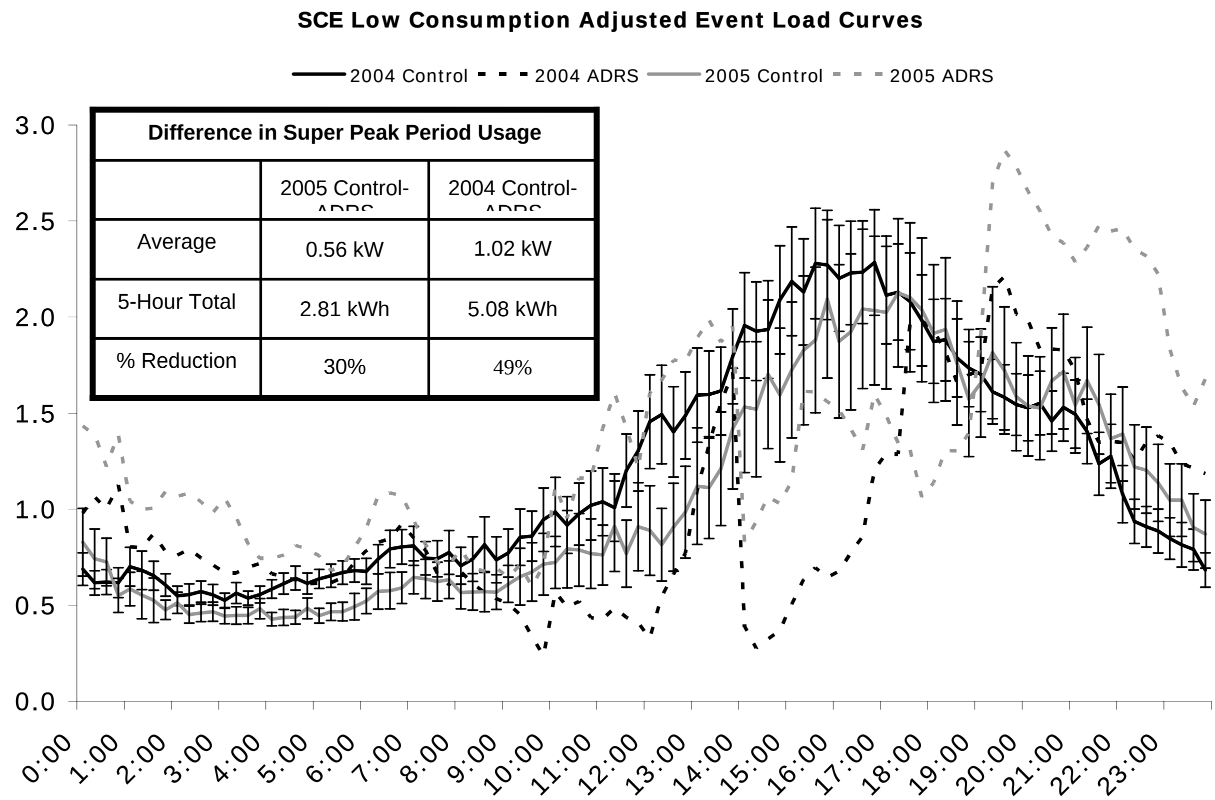


Figure 57 Comparison of 2005 and 2004 SCE low consumption homes: average of non-event weekdays load curves

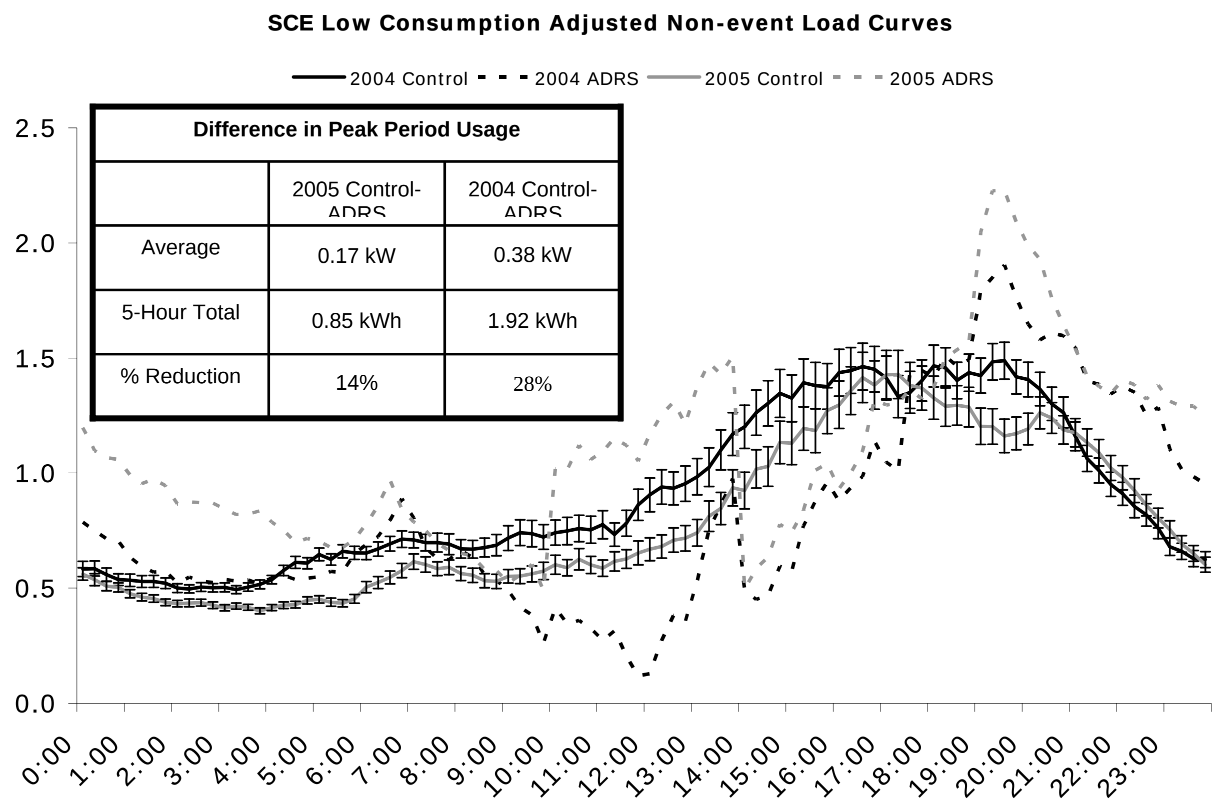
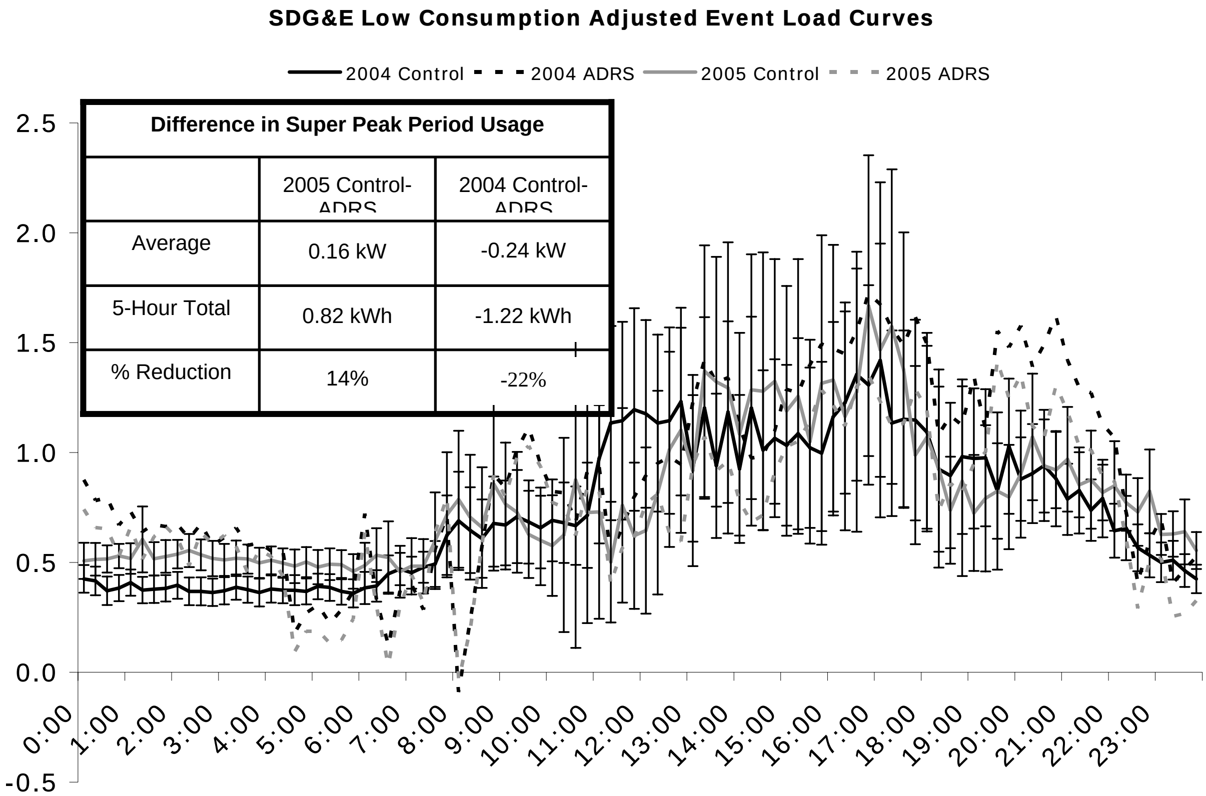


Figure 58: Comparison of 2005 and 2004 SDG&E low consumption homes: average of event day load curves



Difference in Super Peak Period Usage		
	2005 Control-ADRS	2004 Control-ADRS
Average	0.16 kW	-0.24 kW
5-Hour Total	0.82 kWh	-1.22 kWh
% Reduction	14%	-22%

Figure 59: Comparison of 2005 and 2004 SDG&E low consumption homes: average of non-event weekdays load curves