Using Behavioral Science to Maximize Savings on Time-of-Use Rates

Will Sierzchula & Carly Olig, Guidehouse
Renee Guillory, Sarah Noll, & Elizabeth Lawrence, Arizona Public Service (APS)

ABSTRACT

Due to a significant summer peak, Arizona Public Service (APS) is working to proactively manage customer load to shift usage away from peak times. APS runs several direct load management programs to achieve this and is also looking to behavioral science for impacts without direct intervention. APS has paired two behavioral programs, Plan Coach and High Usage Alerts (HUA), with their residential time-of-use (TOU) and demand rates to encourage participants to shift their load to off-peak hours. Plan Coach provides participants feedback about their usage, and tips to optimize it based on their rate plan, via weekly emails. HUA frames messaging using loss aversion to incentivize reducing overall usage or shifting usage to off-peak hours to reduce what could be higher-than-normal usage for the current bill cycle.

But are these behavioral nudges effective? Results show both HUA and Plan Coach are effective stand-alone measures at shifting peak demand with per-customer maximum hour savings ranging from 0.026 kW to 0.064 kW. The measures can also supplement Home Energy Reports (HER) to provide an additional 0.019 kW to 0.026 kW of peak demand reduction on top of HER savings. While impacts varied across waves, this study found Plan Coach to be more effective at reducing demand than HUA. HUA and Plan Coach demand impacts were considerably higher (4% - 13%) than energy impacts (-0.31% - 0.36%); as expected since load management is the primary goal. Attendees of this session will gain ideas for how to leverage behavioral science to further achieve grid management objectives.

Introduction

Due to a significant summer peak, APS is working to proactively manage customer load to shift usage away from peak times while maintaining mutual benefits for the grid and customers. APS is also committed to providing 100% clean, carbon-free energy by 2050 while maintaining reliability and affordability for customers. Load management programs help transition to a clean energy future by enabling more integration of renewables. APS runs several direct load management programs to achieve its clean energy goals and is also looking to behavioral science—and conservation behavior programs—for impacts without direct control. It is important to APS to have different tools and products available to customers, while ensuring a positive customer experience. A key factor to the success of APS's conservation behavior portfolio has been the clear and actionable messaging to customers. APS is focused on delivering an experience that informs, educates and provides a meaningful way to engage with customers to help them understand how to move from simply reducing overall energy usage to reducing energy usage on-peak, while maintaining comfort during Arizona's extreme summers.

Specifically, this paper looks at two APS behavioral programs, Plan Coach and High Usage Alerts (HUA), that are paired with residential time of use (TOU) and demand rates to encourage participants to shift their load to off-peak hours. We provide an overview of behavioral programs designed to save energy or shift demand; detail some of APS's behavioral programs, specifically their Plan Coach and HUA offerings; and finally provide energy and demand impacts from those programs along with APS's Home Energy Report (HER) measure as a comparison.

Our analysis shows both HUA and Plan Coach are effective stand-alone measures at shifting peak demand with per-customer maximum hour savings ranging from 0.026 kW to 0.064 kW. The measures can also supplement HER to provide an additional 0.019 kW to 0.026 kW of peak demand reduction on top of HER savings. We found Plan Coach to be more effective at shifting load than HUA and it produced the added benefit of energy savings when used as a stand-alone measure.

History of Behavioral Programs

For decades, utilities and implementers have designed behavioral programs to generate energy savings by providing residential customers with information about energy use and conservation. Oracle's Opower product has been a pioneer in this space providing participants with HERs that give customers insight into their energy use, including:

- An assessment of how the customer's recent energy use compared to past energy use.
- Tips on how to reduce energy consumption or demand, some of which were tailored to the customer's unique circumstances.
- Information on how their energy use compares to that of customers with similar homes.

Utilities have used HERs for more than a decade to help reduce household energy consumption. These HERs are sent to customers via paper or email and have traditionally been closely tied to monthly bills. HER programs often require a year-long ramp-up period as customers process the communication and adjust their energy-saving behavior accordingly. Afterwards, objective third-party evaluations have shown HERs to save between 0.5% to 3.0% of a customer's whole-home energy usage (Goldman et al, 2020).

With the rise of Advanced Metering Infrastructure (AMI) and variable rate plans, behavioral programs have also evolved as load shaping tools for utilities. They can be used to encourage load control for isolated events such as behavioral demand response (BDR) or to encourage load shifting away from all peak hours. A literature review found BDR events reduce demand during event hours by 1% - 4%, or 2.6%, on average (Guidehouse, 2019).

Variable rate plans, such as TOU rates and demand pricing, allow customers who select these plans to manage their costs by shifting energy use to lower-cost off-peak hours and use less energy during the higher-cost on-peak hours, which reflects constraints on the grid and the amount it costs utilities to generate power during those times. These plans, along with behavioral load shifting programs, help to align utility and customer incentives. Customers can save more by reducing demand during peak hours. For example, households can shift loads of laundry or electric vehicle charging from evening on-peak hours to later in the night. 6-10 p.m. This need to reduce demand during peak hours is heightened in parts of the country that experience high temperatures and the subsequent need to provide the power for air conditioners during summer to make homes more comfortable. With APS, the load-shifting strategy also serves to encourage using energy when more solar resources are available. Broadly speaking, successful load-shifting also fosters the integration of more renewable resources, which ties behavioral programs to critical enterprise goals on decarbonization.

APS' Behavioral Load Management

APS has had TOU rates in place for over 40 years. However, to get the most value from their time variant rate APS customers needed help understanding how to shift their load a. APS has focused on creating a seamless customer experience, while increasing customer satisfaction. With Plan Coach and HUA, APS created an experience that informs and educates participants to help them understand how to move from reducing energy usage generally to reducing energy usage *on-peak*. Because other conservation behavior programs had demonstrable success helping customers save energy, APS also wanted to provide customers on TOU rates with an opportunity for more savings through behavioral load shaping. APS has two main TOU service plans, illustrated in Figure 1. Both have on-peak hours from 4–7 p.m. on non-holiday weekdays and one of them also has a demand charge for the highest hour of usage during that time. Participants in APS's Conservation Behavior program can also be on a frozen TOU service plan that has on-peak hours from 3–8 p.m.. As described below, APS is also using behavioral measures to help customers optimize their usage on these dynamic rates.

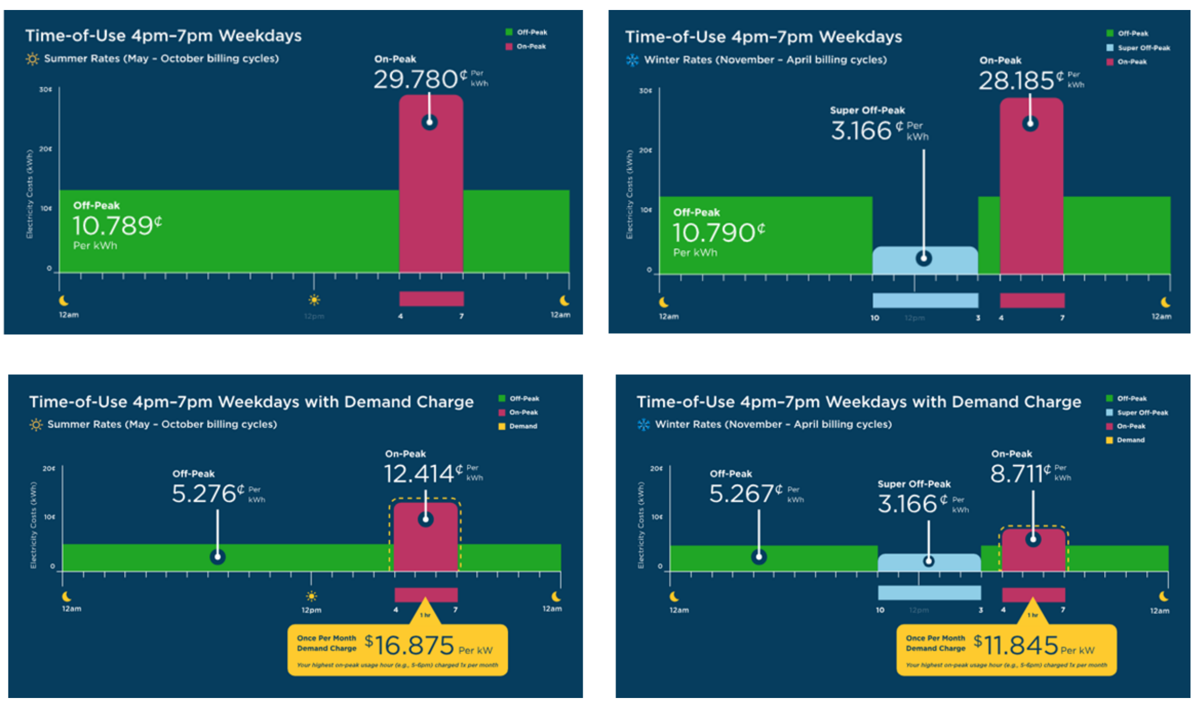

Figure 1. APS TOU service plan illustration. APS rates have since changed, but Figure 1 shows the rates in effect during program year 2023. *Source*: APS.

APS runs an extensive residential Conservation Behavior program with Oracle as the implementer. The program first launched to a group of 70,000 customers as an HER measure in 2011 with a goal of driving energy efficiency savings. In 2023, APS sent HERs to over 500,000 participants and achieved energy savings of 81.28 kWh/participant (0.7%) for a total of 47 GWh saved. APS is actively exploring ways to expand the program to customers who have historically been ineligible, making the program's reach even greater.

Over the years, APS has expanded the reach of the Conservation Behavior program beyond energy efficiency savings to also target peak demand savings and load management. Plan Coach was launched in 2020 and provides participants with information to assist them in optimizing their energy usage for their service plan. Participants on TOU and demand rates are coached to shift usage away from their on-peak rate hours. In the summer of 2022, APS added Energy Saving Days to the program, a BDR measure that, in the 2023 season, alerted over 327,000 participants to the need to save energy during five events and resulted in maximum hour savings of seven MW. In November of 2022 HUAs were added to the program to alert participants when they are using more energy than usual and assist them in reducing their energy in a cost-effective way, again guiding participants on TOU and demand rates to shift away from peak usage. In 2023, there were 370,000 participants across all plan coach waves: 120,000 customers on the fixed charge rate and 250,000 customers on the APS Time-of-Use 4-7 p.m. Weekdays and Time-of-Use 4-7 p.m. Weekdays with Demand rates (most of whom also receive HERs) and about 90,000 HUAs were sent to eligible customers. Figure 2 provides a graphical overview of APS's behavioral programs.

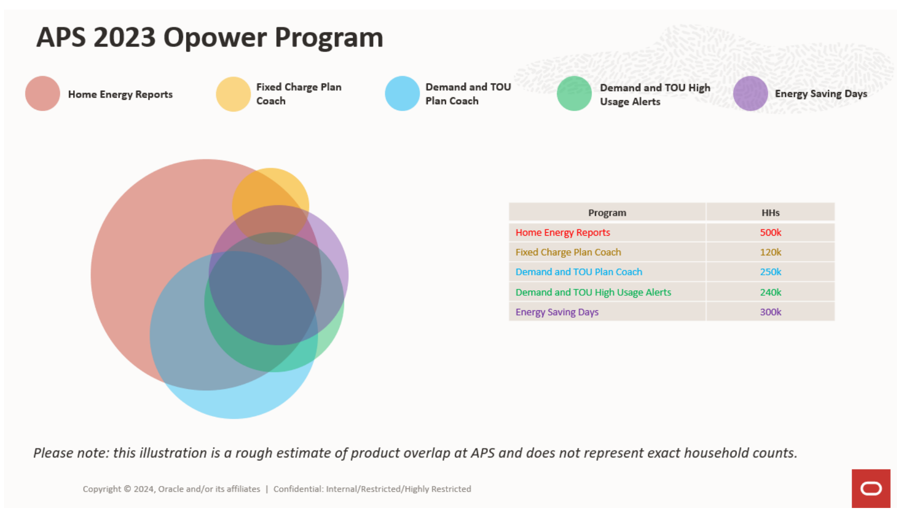

Figure 2. Graphic display of APS behavioral programs.

The next section provides more detail on Plan Coach and HUA, followed by sections on the testing setup (i.e., customer service plan and measure evaluated), methods used to estimate savings, the results, and a conclusion.

Plan Coach

Plan Coach provides participants with a weekly email about their on-peak usage and advice on how to reduce or shift demand during these hours. Plan Coach messaging is sent to almost all participants who receive HERs in addition to the waves described specifically for testing in this paper. Figure 3 below illustrates an example of the communication Oracle sends to Plan Coach participants to induce behavioral load-shifting changes.

APS, Oracle, and Guidehouse developed two waves to measure the incremental impacts of Plan Coach on top of HERs for participants on demand rates. In these waves, all participants are receiving HERs, so any impacts would come solely from the incremental gains associated with the Plan Coach offering. One wave is designed to estimate the incremental impact of just Plan Coach, while the other includes the combined impacts of Plan Coach and HUA. The program also includes two waves that get Plan Coach without also receiving HERs: one that receives Plan Coach alone, and one that also receives HUA.

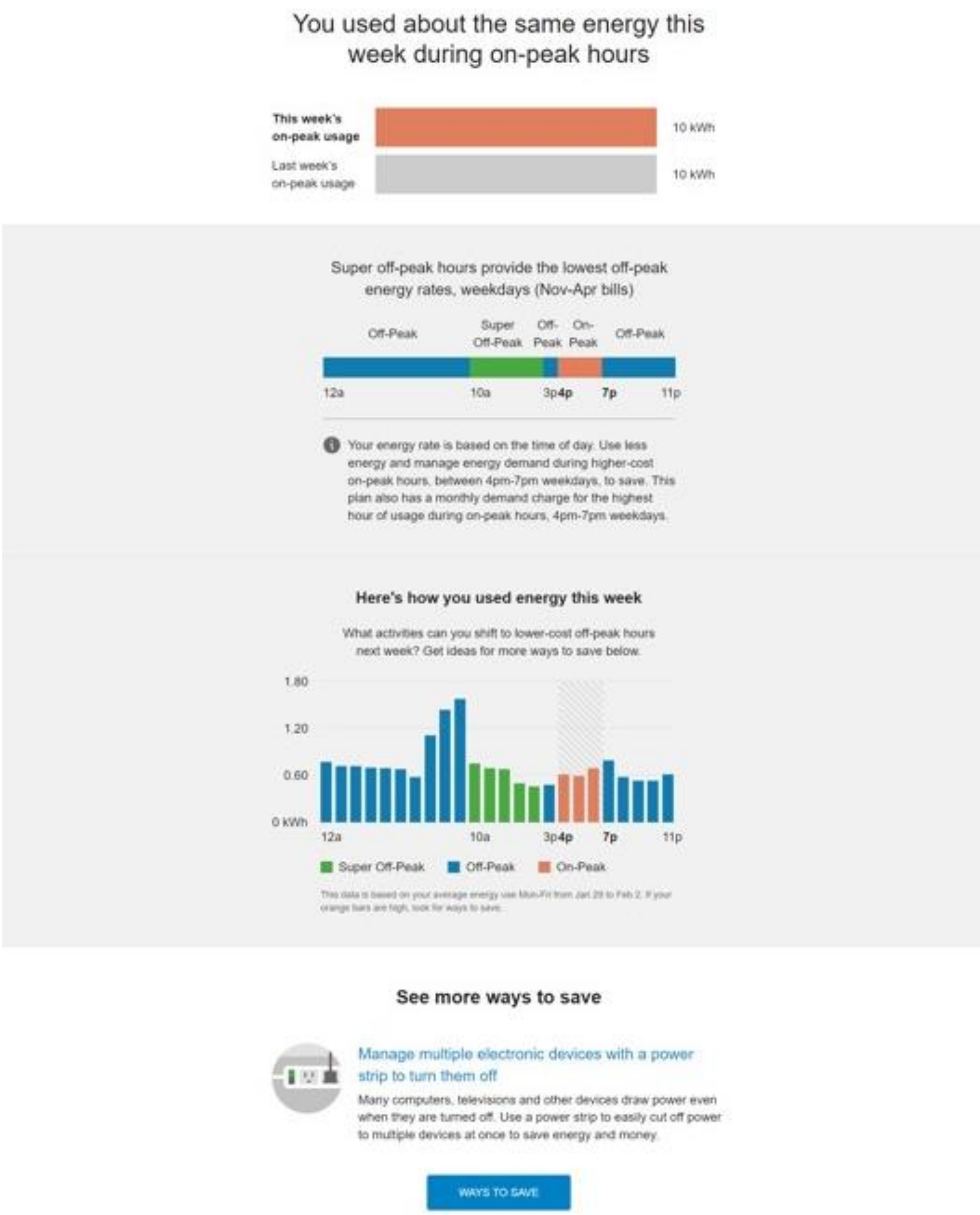

Figure 3. Plan Coach communication example. *Source*: Oracle.

High Usage Alerts

HUA frames messaging using loss aversion to incentivize reducing overall usage or shifting from expensive to cheap hours to reduce what could be a higher-than-normal usage for the current bill cycle. HUAs are triggered communications and will only be sent when a customer's usage is projected to be 30% higher as compared to the same time in the previous year. In 2023, fewer than 10% of HUA-eligible customers actually received a HUA. APS, in collaboration with Oracle and Guidehouse, set up two experimental waves to estimate HUA impacts incremental to HERs and two to measure HUA impacts without HERs. These waves are summarized in the next section.

Testing Setup

Guidehouse, APS, and Oracle worked together on a program setup to allow Guidehouse to evaluate energy and demand savings driven by Plan Coach and HUA both when provided in conjunction with HERs and in isolation. To this end, APS has three program waves, all for demand service plan customers, which do not receive HERs: one receives HUA only, one receives Plan Coach only, and one receives both. In addition, there are three groups where all participants receive HERs, but the treatment group also receives Plan Coach and/or HUA.[1] This experimental setup allows Guidehouse to measure the incremental impact of Plan Coach and/or HUA on top of HER savings as well as these measures in isolation from HERs. Table 1 below summarizes this program setup.

Table 1. APS behavioral program wave names and impact areas

Wave	Launch	Service plan	# of participants	Plan Coach	HUA
Non-HER HUA	12/2022	Demand	10,283		x
Non-HER PC	12/2022	Demand	9,498	x	
Non-HER HUA and PC	12/2022	Demand	26,338	x	x
Incremental HUA †	11/2022	TOU	121,780		x
Incremental PC †	12/2022	Demand	10,498	x	
Incremental HUA and PC †	12/2022	Demand	28,989	x	x

†Wave impacts are incremental to HER (i.e., both participants and controls receive HER).

Methods

Each of APS's residential Conservation Behavior program measures rely on a randomized controlled trial (RCT) to estimate savings wherein customers targeted for the measure are randomly divided between a treatment group who receives the measure and a control group who does not. RCT is the gold standard to estimate program-related impacts because randomization ensures an unbiased estimate of savings (Stewart and Todd, 2020). Guidehouse relied on the RCT design to estimate savings for each of the Plan Coach and HUA waves described in the section above via regression analysis.

[1] See the methods section below for more information about the randomization.

Randomization

To test that behavioral waves were consistent with an RCT, Guidehouse compared treatment and control usage for each month during the pre-program period. If the allocation of households across participants and controls is truly random, the two groups should have the same distribution of energy usage during the twelve months prior to receiving the program intervention. Guidehouse conducted variance tests and t-tests comparing participant and control usage for each month of the pre-period and found that mean usage was not statistically different. As an additional check, we performed a regression analysis in which average daily usage in the pre-program period was a function of monthly binary variables and a binary participation variable which showed participation did not impact usage. The analysis validated randomization between participants and controls for each of the waves described in Table 1.

Data Cleaning

In preparation for impact analysis, Guidehouse combined and cleaned AMI data provided by APS, monthly billing data provided by the implementer, and a list of participant and control customers from the implementer. The dataset included 207,378 treatment customers and 65,193 controls. Data during the twelve-month pre-period for each wave and throughout 2023 was used in the regression analysis for each of the lagged dependent variable (LDV) models described below. Examples of data cleaning steps include removing outliers and negative usage values.

Regression Modeling

The regression model measured the difference in average energy or demand between participants and customers randomly assigned to the control group.

Demand Lagged Dependent Variable Model

Guidehouse's demand LDV model controls for non-treatment differences in demand between treatment and control customers using lagged use as an explanatory variable. The model frames use in hour *t* of the evaluation period as a function of both the treatment variable and average use in the same hour of the same calendar month of the pre-program period. The underlying logic is that systematic differences between control and treatment customers will be reflected in differences in their past use, which is highly correlated with their current use. The regression model was run monthly by wave to measure the difference in average demand between participants and controls. β_1 is the coefficient that indicates per-customer hourly demand impacts.

$$Hourly.usage_{kt} = \beta_1 Treatment_{kt} + \beta_2 Avg.pre.same.month_{kt} + \varepsilon_{kt}$$

Where:

- $Hourly.usage_{kt}$: indicates the average consumption for household *k* in hour *t* during the program period
- $Treatment_{kt}$: indicates whether household *k* was a treatment (i.e., received HUA or Plan Coach communications) or control customer

- $Avg.pre.same.month_{kt}$: indicates the average consumption for household k in hour *t* in the same calendar month in the year prior to program treatment
- ε_{kt}: cluster-robust standard error for household *k* in hour *t*; cluster-robust errors account for heteroskedasticity and autocorrelation at the household level

Energy Lagged Dependent Variable Model

Guidehouse's energy LDV model similarly controls for non-treatment differences in energy use between treatment and control customers using lagged energy use as an explanatory variable. Formally, the model is shown below. Like the model above, β_1 is the coefficient that indicates per-customer monthly energy impacts.

$$ADU_{km} = \beta 1 Treatment_k + \sum_j \beta_{2j} Month_{jm} + \sum_j \beta_{3j} Month_{jm} * ADUlag_{km} + \varepsilon_k$$

Where:

- ADU_{km}: is average daily consumption of energy by household *k* in bill period *m*
- $Treatment_k$: is a binary variable taking a value of 0 if household *k* is assigned to the control group, and 1 if assigned to the treatment group
- $Month_{jm}$: is a set of monthly fixed effects taking a value of 1 when j=m and 0 otherwise
- $ADUlag_{km}$: is household *k's* energy use in the same calendar month of the pre-program year as the calendar month of month *m*
- ε_k: is the cluster-robust error term for household *k*; cluster-robust errors account for heteroskedasticity and autocorrelation at the household level

Results

Guidehouse estimated demand and energy impacts for Plan Coach and HUA either as stand-alone measures or layered on top of HER. This section provides tables and graphs of our results for each wave along with whole home impacts. More discussion of the results is provided in the conclusion section.

Demand Results

For all the studied waves, the single highest hour of demand reduction occurred during the on-peak rate hours (4–7 p.m.) of the summer months (June–September). The studied waves showed a wide range of demand impact in the single highest hour with the incremental HUA wave only reducing peak demand by 0.019 kW, while the non-HER Plan Coach wave reduced kW by three times as much. Figure 4 below shows the maximum demand impact for the six waves in this study along with 90% confidence intervals.

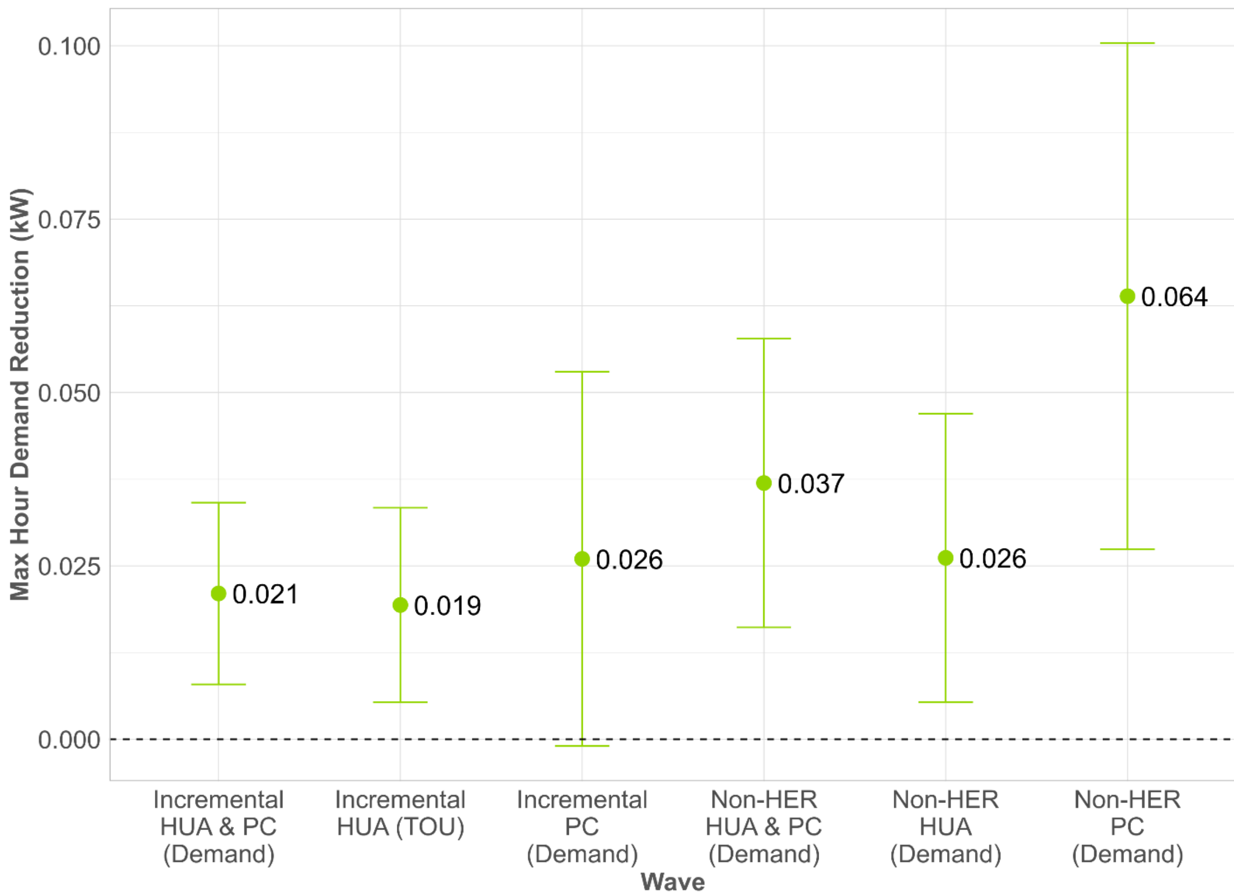

Figure 4. Max hour demand impacts by wave with 90% confidence intervals.

Table 2 shows the average demand reduction during the on-peak rate hours of the summer months (4–7p.m. in June–September), in addition to the maximum shown in Figure 3. The table also converts the maximum hour's savings into a percentage by dividing by baseline usage in the same hour. The non-HER waves tended to exhibit higher demand impacts, both average and maximum, which is expected because the incremental waves already have some demand reduction incorporated from HER.

Table 2. Plan Coach and HUA demand impacts

Wave	Service plan	Average demand reduction	Max hour demand reduction	Max hour demand baseline	Max hour % whole home reduction
Non-HER HUA and PC	Demand	0.023 kW	0.037 kW	0.289 kW	13%
Non-HER PC	Demand	0.055 kW	0.064 kW	0.548 kW	12%
Non-HER HUA	Demand	0.003 kW	0.026 kW	0.319 kW	8%
Incremental HUA	TOU	0.009 kW	0.019 kW	0.507 kW	4%
Incremental PC	Demand	0.008 kW	0.026 kW	0.459 kW	6%
Incremental HUA and PC	Demand	0.013 kW	0.021 kW	0.449 kW	5%

Figure 5 shows average summer hourly demand impact load shapes for each wave and the grey shading indicates APS's on-peak rate hours of 4-7 p.m. We chose to show impacts during summer because those are the months with the highest demand, primarily due to air conditioning use in Arizona, and the highest savings. Instances where the blue line goes below zero (the dashed line) were hours when the treatment customers who received HUA or Plan Coach communications had lower demand than the control customers (after controlling for other factors), i.e., savings. For example, the Non-HER Plan Coach wave for demand rate participants showed consistent demand reduction across all hours with the highest impacts during on-peak hours. Waves that include Plan Coach have distinctive increases in load reduction during on-peak rate hours, while the HUA-only waves do not show a similar pattern.

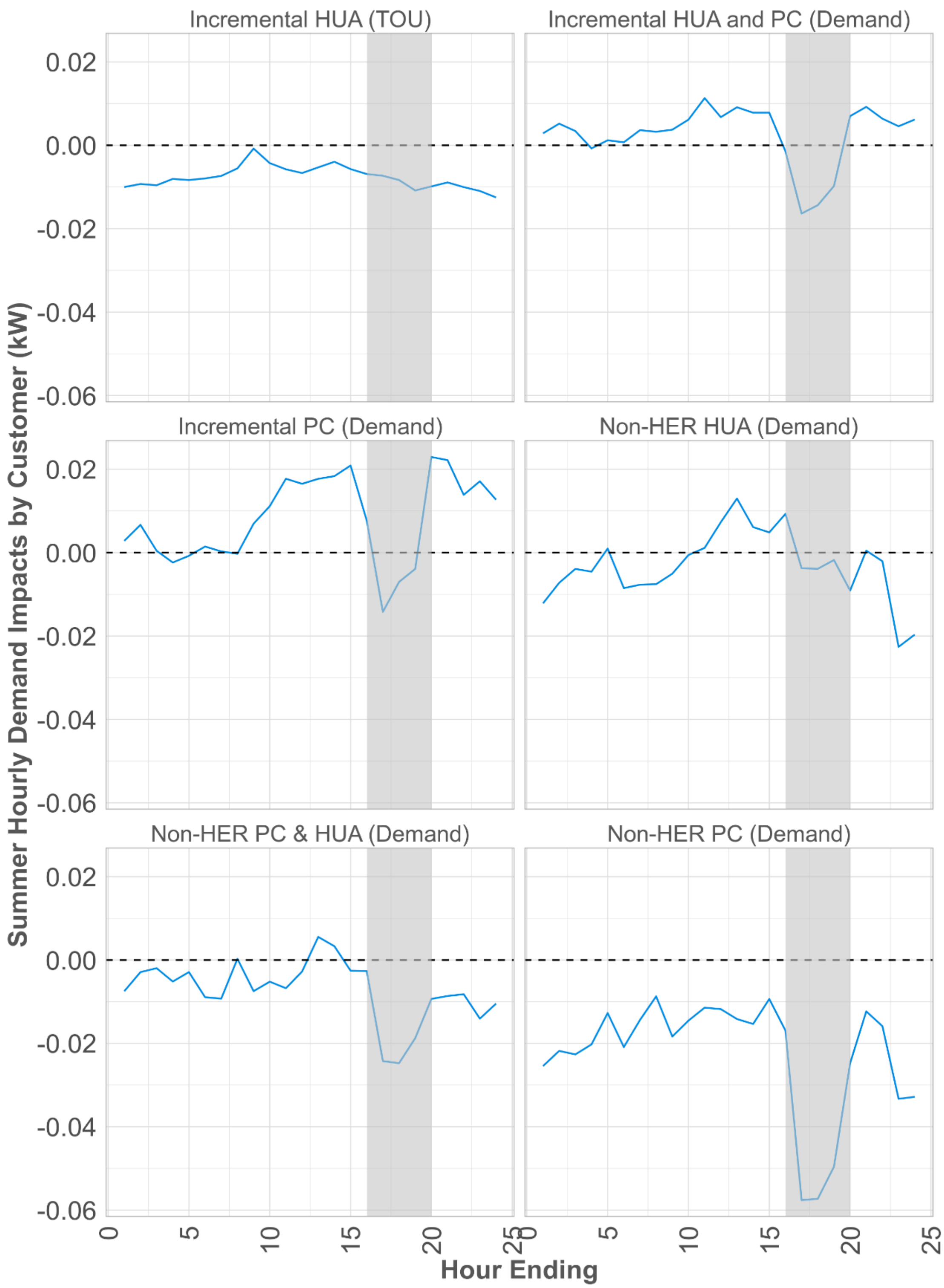

Figure 5. Load shifting impacts during summer months (June – September) by wave with shaded on-peak rate hours. Values below 0 indicate hourly demand reduction, due to the measure.

Energy Results

Table 3 shows per customer and whole home energy savings for the waves evaluated, along with HER savings, without Plan Coach or HUA), for comparison. Only one wave, non-HER PC, had savings that were statistically different from zero at the 90% confidence level (illustrated in Figure 5 below), which is important to keep in mind while reviewing the table and this text which is commenting on the point estimates. Across the three test waves without HER, energy savings averaged 0.1%. For comparison, HER waves, without Plan Coach or HUA, average 0.67% energy savings. Any energy savings for these waves are a bonus, as the primary goal of these measures is peak demand reduction. The wave with only Plan Coach had the highest energy savings of the non-HER test set, while the one with Plan Coach and HUA was a bit lower but not statistically different with 90% confidence. The HUA trial wave without HER showed dissaving (the point estimate was negative though not statistically significant). These results suggest that in the absence of HER, Plan Coach drives more energy savings than HUA. One explanation for this result is Plan Coach provides consistent monthly communication while HUA only reaches out when triggered. For the three test waves incremental to HER, HUA for participants on TOU rates had the highest energy savings among the waves analyzed at 50.47 kWh, while the remaining incremental waves showed dissavings (though again not statistically significant). Notably, HUAs demonstrated peak reductions in usage even though fewer than 10% of HUA-eligible customers actually received a HUA.

Table 3. Plan Coach and HUA energy impacts

Wave name	Service plan	Annual energy savings (per participant)	% whole home savings
Home Energy Report	TOU & Demand	81.28 kWh	0.67%
Non-HER HUA	Demand	-34.15 kWh	-0.18%
Non-HER PC	Demand	36.55 kWh	0.19%
Non-HER HUA and PC	Demand	23.96 kWh	0.12%
Incremental HUA	TOU	50.47 kWh	0.36%
Incremental PC	Demand	-61.17 kWh	-0.31%
Incremental HUA and PC	Demand	-19.12 kWh	-0.10%

Figure 6 illustrates these values along with 90% confidence intervals. As indicated in Table 3, only three waves saw positive point estimates and only one (incremental HUA for TOU rate participants) showed impacts where the difference from zero was statistically significant at the 90% confidence level. Across studied waves, whole home energy savings were less than 0.5%. Due to this low savings value, utilities that need results to be statistically significant should deploy waves with tens of thousands of participants to achieve the appropriate statistical power.

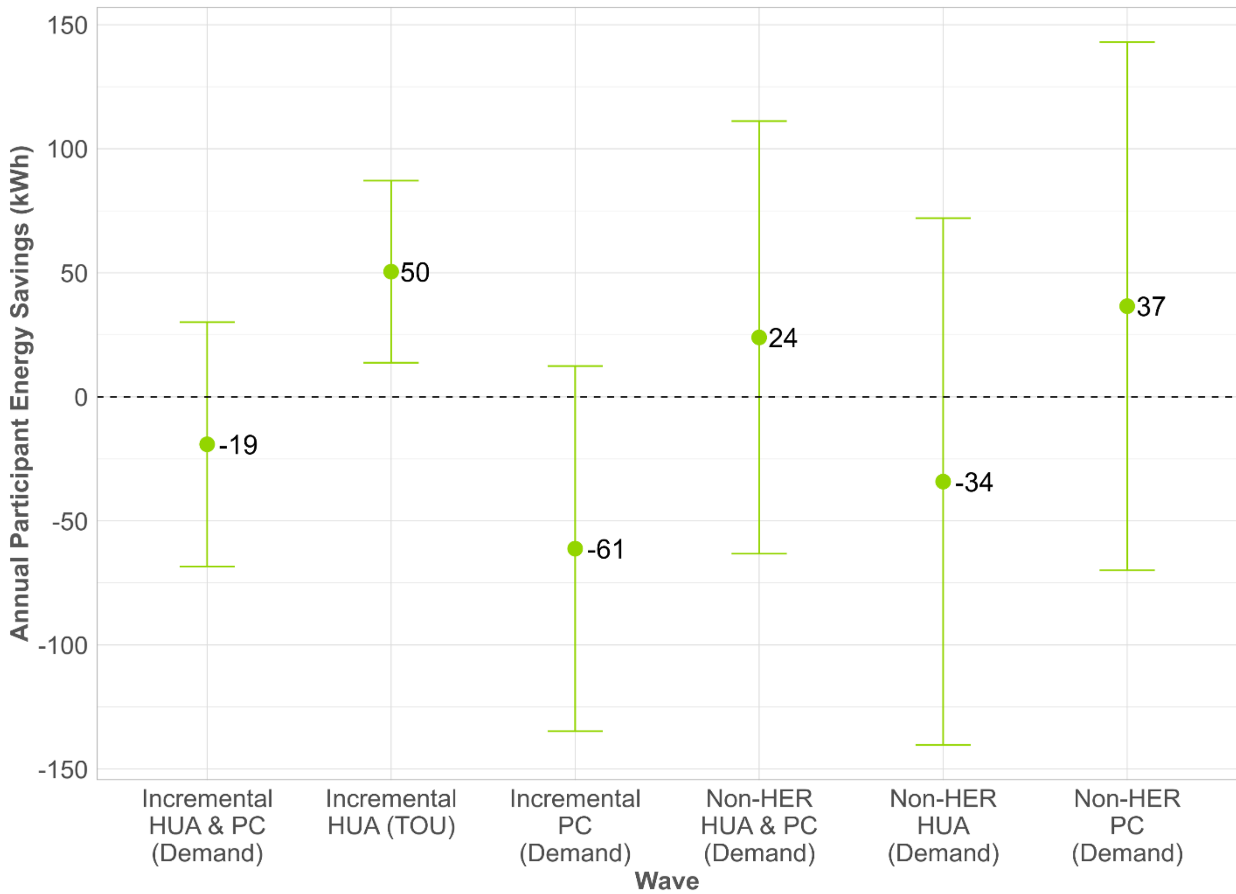

Figure 6. Energy impacts by wave with 90% confidence intervals.

Energy savings by wave were not consistent throughout the year. Behavioral program impacts often experience ramp-up where savings increase over the first year, which is what we saw for all three of the non-HER waves. Figure 7 illustrates this, showing incremental monthly savings for the Non-HER Demand Plan Coach wave, which increased as the program progressed in 2023. The other two non-HER waves followed a similar trend.

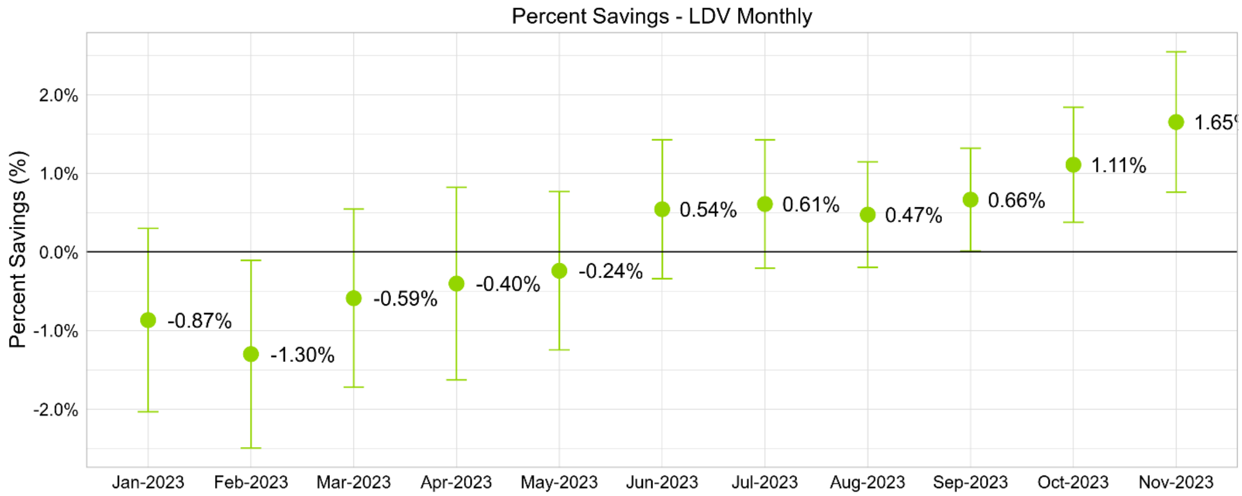

Figure 7. Non-HER Plan Coach demand whole home savings.

Conclusion

Guidehouse's analysis shows both HUA and Plan Coach are effective stand-alone measures at shifting peak demand with per-customer maximum hour savings ranging from 0.026 kW to 0.064 kW. The measures can also supplement HER to provide an additional 0.019 kW to 0.026 kW of maximum hour peak demand reduction on top of HER savings. While impacts varied across waves, this study found Plan Coach to be more effective at reducing demand than HUA. This is illustrated in Figure 4 where we see a big jump in demand reduction during the peak hours for waves with Plan Coach, but not for those with HUA alone. This is likely in-part, since Plan Coach provides consistent monthly communication, while HUA only reaches out when triggered.

On a whole home energy basis, HUA and Plan Coach demand impacts were considerably higher (4% - 13%) than energy impacts (-0.31% - 0.36%), as expected since load management is the primary goal of these measures. HUA and Plan Coach showed mixed results for saving energy.[2] This is illustrated in Figure 5 where we can see that only 3 of the 6 waves had positive point estimates for energy savings, and only one was statistically significant with 90% confidence.

There are some factors which may explain why the incremental HUA TOU wave had such high energy savings, compared to the other waves. The percent of participants who received HUA communication was 30% higher for the TOU wave than the demand waves. Communication is variable because HUA are only sent when a customer's usage is projected to be a certain proportion higher as compared to the same time of the previous year. The TOU rate participants received more communication, and it is also possible that TOU participants are more responsive to HUA communications because they may have more opportunity to change their usage and save money on their bill, as compared to demand rate participants who may have already had their peak hour of usage, which drives their bill amount, by the time they receive the communication and cannot go back in time to change it.

---

[2] Stand-alone Plan Coach waves did show consistent energy savings while the other waves did not.

An important caveat for these first year of energy savings is that behavioral programs often involve a ramp-up period where savings start low, increase over time, and then stabilize. Guidehouse observed a similar pattern of increasing savings through the first year for all non-HER waves in this analysis. For example, energy savings in the non-HER Plan Coach wave increased from -0.87% in January to 1.65% in November. A second year of analysis could confirm whether monthly savings stabilize at a higher level as is commonly observed in multi-year HER program evaluations.

As utilities seek to reduce peak demand and save energy, there are a host of options for them to consider. This analysis provides evidence that behavioral measures, specifically Plan Coach and HUA, can drive demand savings when targeted around TOU rates. These results may be helpful to other utilities as they decide which measures to include in their demand-side management portfolio.

References

Goldman, Charles A., Ian Hoffman, Sean Murphy, Natalie Mims Frick, Greg Leventis, and Lisa Schwartz. 2020. Energies 2020. *The Cost of Saving Electricity: A Multi-Program Cost Curve for Programs Funded by U.S. Utility Customers*. https://emp.lbl.gov/publications/cost-saving-electricity-multi-program

Guidehouse, 2019. *2019 National Grid Behavioral Demand Response Evaluation Findings*. Prepared for National Grid. https://ma-eeac.org/wp-content/uploads/NG-BDR-2019-Findings-Report-FINAL-2019-04-28.pdf

Stewart, James and Annika Todd. 2020. *Chapter 17: Residential Behavior Evaluation Protocol, The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures: September 2011 – August 2020*. Golden, CO: National Renewable Energy Laboratory. NREL/SR-7A40-77435. https://www.nrel.gov/docs/fy21osti/77435.pdf.