Supercharging Demand Response Performance with Residential Batteries

Carly Olig, Ari Kornelis, Gabrielle Bethke, Guidehouse
Iain Shovelin, Jason Delaney and Elizabeth Lawrence, Arizona Public Service (APS)

ABSTRACT

Residential batteries can deliver more power than traditional residential demand response (DR) resources such as thermostats but getting the most value out of these new grid resources requires careful planning. Arizona Public Service (APS) is piloting behind-the-meter residential battery storage DR with customers on time-of-use (TOU) rates with solar photovoltaic systems.

By the end of 2023, APS had nearly 200 customers with battery systems enrolled for DR events. Customers received an incentive on the purchase of a battery system in exchange for sharing performance data with the utility and received an additional incentive to enroll in battery dispatch management solely during up to 100 DR events per year. This incentive is in addition to savings customers can receive by using their battery to manage load in response to TOU rates. Events were called between 4-9 p.m. Peak TOU hours are 4-7p.m. for most customers.

This paper shares findings from a 2023 evaluation of the pilot and covers several topics, including: (1) what to expect from events, (2) coordinating event dispatch with dispatch from TOU rates, and (3) challenges by battery partner. On average, we estimated weekday events had 1.36 kilowatts (kW) of additional hourly discharge compared to baseline discharge in the absence of an event. These results show storage systems do have charge left after the on-peak period to continue discharging during events. Challenges across partners included event failures; pre-event management, which may affect participant bills; reaching the 20% state-of-charge reserve capacity; and systems discharging less than their theoretical limit.

Introduction

Arizona Public Service (APS) launched a Residential Battery Pilot in October 2021. The pilot is intended to help APS learn about battery performance and how batteries may create value for customers through improved management of energy use at their residence and help reduce stress on the electric grid. A full description of the program occurs in the next section. In brief, participants received an upfront incentive when purchasing their battery by agreeing to either share their battery system performance data with APS (Data Only – DO) or allow APS to manage up to 80% of their battery's capacity during demand response (DR) events[1] (Data and Management – D+M). This program is unique among battery pilots as the majority of participants are also on the APS Time-of-Use 4-7 p.m. Weekdays and Time-of-Use 4-7 p.m. Weekdays with Demand Charge rates giving them an economic incentive to program their batteries for discharge during their on-peak rate hours. These rates are described in more detail in the next section.

Due to a significant summer afternoon load peak, APS is working to proactively 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

[1] In addition to sharing their battery performance data.

reliability and affordability for customers. The Residential Battery Pilot contributes to APS's load management needs in two ways:

(1) because of the TOU rates, all participants have an economic incentive to program their batteries for discharge during their on-peak rate hours. This reduces their need to pull power from the grid during these times.
(2) D+M participants have their battery managed by APS to discharge during DR events, further reducing grid load during events.

This paper focuses on the D+M portion of the program and response to the program's DR events. For more information on the non-event program impacts see Shovelin and Kornelis (2023). That research found, outside events when batteries are not managed by APS, on average, customers reduced their usage pulled from the grid by about 50% during the TOU peak period from 4-7 p.m. on weekdays. This effect was driven by 50-75% of systems that were discharging during the TOU peak period; the rest of systems were typically idle (neither charging nor discharging) during this period with only a very small portion meaningfully charging, as 4-7p.m. is after peak solar production when battery systems do most of their charging. On non-event days the system state-of-charge (SOC) after the TOU period averaged approximately 50% across all customers (DO and D+M). The DR events in the D+M portion of the program take advantage of battery capacity by customers who are not discharging during the TOU peak hours and any capacity left for those customers who do discharge.[2] The DR events can also occur on the weekend when there is no TOU peak period, though some customers do still have their batteries programmed to discharge between 4-7 p.m. even on the weekend.

The APS program is unique among most battery DR programs in combining DR events with TOU rates. Because many customers discharge their battery daily[3] in response to their TOU rate, the baseline for DR events in this evaluation is not no battery where the event impact is the entire discharge. Instead, the baseline we chose is the counterfactual of what amount of discharge the customer would have had in the absence of the event to appropriately separate the impact of the event from what customers are just doing daily.

Program Overview

Through this pilot, customers can receive incentives when buying a new battery from participating vendors (Enphase, SolarEdge, Sunrun, and Tesla) if they enroll in one of the following two options:

- Option 1: Data Only (DO) – Customers receive $500 per kilowatt (kW) of battery system discharge capacity installed (up to $2,500 per home) if they agree to share their battery system performance data with APS. There is no APS or implementer management of the battery charge and discharge in this option.
- Option 2: Data and Battery Management (D+M) – Customers receive the same incentive as option 1 plus an additional $1,250 (for a maximum of $3,750 per home) if, in addition to sharing their battery performance data, they also allow APS to manage up to 80% of their battery's capacity during up to 100 events per year. Events can last one to four hours

[2] On weekdays DR events occur between 6 p.m. and 9 p.m. overlapping with the end of and following the TOU peak hours.
[3] This is under the customer's own choosing, not managed by APS.

and be called between 6-9 p.m. on weekdays and 9 a.m. - 9 p.m. on weekends and holidays. Events are run through EnergyHub, the program aggregator.

The batteries in this pilot can, but do not have to, be paired with a new or existing solar photovoltaic (PV) system. Pilot participants must be on a time-of-use (TOU) service plan or a grandfathered service plan with a solar rate rider.[4] APS has two main TOU service plans, Time-of-Use 4-7 p.m. Weekdays and Time-of-Use 4-7 p.m. Weekdays with Demand Charge, illustrated in Figure 1: both have on-peak hours from 4-7 p.m. on non-holidays weekdays and one also has a demand charge for the highest hour of usage during that time. Participants in the pilot can also be on a plan that has on-peak hours from 3-8 p.m. and a demand charge for the highest hour of usage during that time[5] or on a handful of grandfathered solar rate plans. Weekday pilot events often overlapped with the on-peak hours for participant's service plans and most continued after on-peak hours to help mitigate snapback on the grid from the end of the on-peak hours.

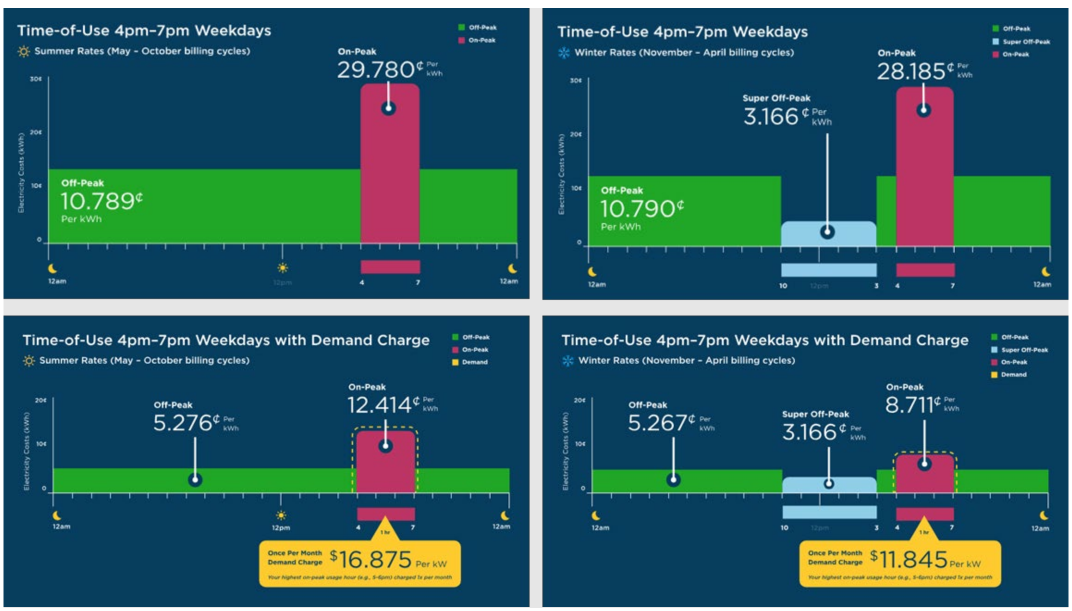

Figure 1. APS TOU service plan illustration. APS rates have since changed, but Figure 1 shows the rates in effect during program year 2023. *Source*: APS.

This paper focuses on the D+M portion of the program and response to the program's DR events.[6] Though this work focuses on D+M and event response, we use the DO participants to help build the baseline for what D+M participants would have done with their batteries in the absence of an event (see the Data and Methods section for more detail).

[4] A solar rate rider defines the terms by which customers who have solar generation at their home are compensated for solar production.
[5] Fewer customers are on this plan as it has been "frozen" meaning customers already on it can stay on it but no new customers can join this plan.
[6] See Shovelin and Kornelis (2023) for more information on how customers are charging and discharging their batteries on a daily basis. A summary of these results appears in the introduction section of this paper.

As of September 2023, 629 storage systems[7] in the pilot had received permission to operate (PTO) from the authority having jurisdiction and the utility. Of these 629 systems, 182 (29%) were enrolled in the D+M portion of the pilot. Figure 2 summarizes systems with PTO in the D+M portion of the pilot by battery partner and service plan. [8] We also found that all systems operating in the pilot as of September 2023 are paired with solar PV.

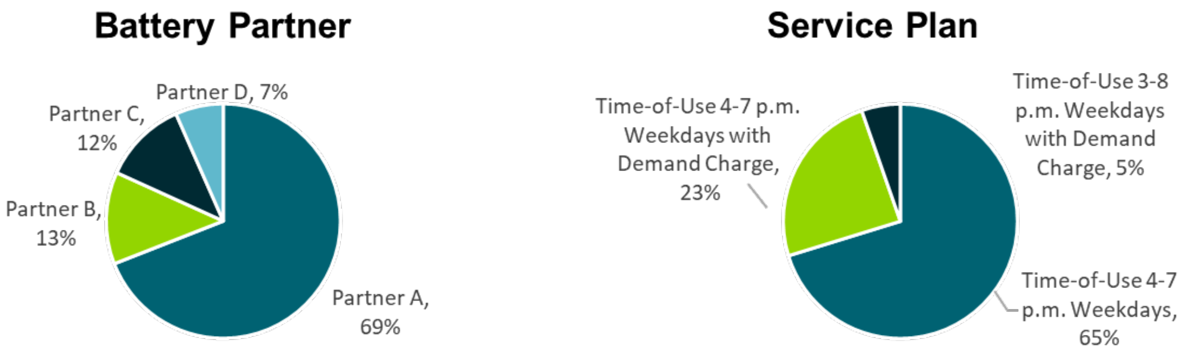

Figure 2. D+M enrollment summary. *Source*: Guidehouse analysis.

While a handful of events were called in 2022, this paper focuses on events called between July and September of 2023 and therefore only reflects results for events called during summer. Most of the 2023 events had nearly all of the 182 D+M systems participating.

As of October 2023, the end of the analysis period, the pilot had approximately five megawatts (MW) of charge/discharge capacity enrolled, with approximately 1.3 MW in D+M. Theoretically, 80% of the D+M energy storage capacity (2.7 MWh) is available for managed discharge during events.[9] Table 1 summarizes the energy storage and charge/discharge capacity of systems in the pilot. Partner A systems are the largest, averaging 23.3 kWh and 8.7 kW. This is about one-third larger than the other partners' systems which average 14.9 kWh and 6.27 kW.

Table 1. Average system size

	Energy Storage Capacity (kWh)	Charge/Discharge Capacity (kW)
Total	13,114	4,994
Data Only	9,764	3,696
Data + Management	3,376	1,298
Avg. per system	30.0	8.0
By Battery Partner (Avg/Sys)		
Partner A	23.3	8.7
Partner B	15.2	6.5
Partner C	16.1	6.1
Partner D	12.2	6.3

[7] Throughout this paper, we will refer to results at the storage system level which we define as one or more batteries tracked in one set of telemetry data. There is an average of 1.6 batteries per storage system.
[8] Partners have been anonymized throughout this paper and we deliberately use a different anonymization in Figure 2 and Table 1 compared to the rest of the report to best preserve that anonymity.
[9] This would be if all systems were fully charged at the beginning of an event and then discharged down to the 20% reserve margin during the event.

	Energy Storage Capacity (kWh)	Charge/Discharge Capacity (kW)
By Program (Avg/Sys)		
Data Only	21.8	8.3
Data + Management	18.7	7.2

Source: Guidehouse analysis.

Throughout July, August, and September 2023, APS called 29 events. There were two-to-four events per week, as shown in Figure 4, with events called on three consecutive days most weeks. Events were two-to-three hours in length and were called between 4 p.m. and 9 p.m. Peak TOU hours are 4-7 pm for most customers. This means some events overlapped with the on-peak hours for participant's service plans and most continued after on-peak hours to help mitigate snapback on the grid from the end of the on-peak hours. Events were also called on weekends when APS does not have on-peak hours but does experience high loads (APS, 2023).

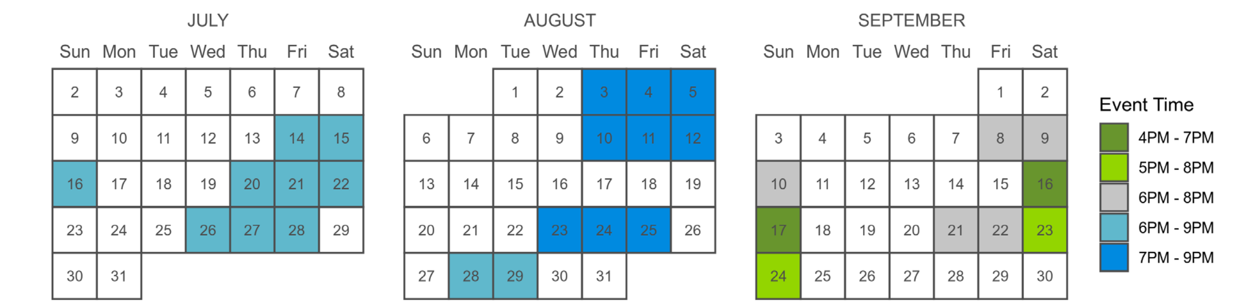

Figure 4. Event summary. *Source*: Guidehouse analysis.

Data and Methods

Data and Data Cleaning

To estimate event impacts Guidehouse received three types of data:

1. Enrollment data – covering customer information such as enrollment date, PTO date, power and energy rating of the storage system,
2. Event data – showing dates and times of the events, and
3. Telemetry data – 15-minute interval data with battery charge, discharge, SOC, and stored energy, as well as site demand and solar production

To help other utilities considering launching similar battery storage programs in their planning, Guidehouse has summarized the data challenges we faced into four categories.

1. Reconciliation of Participant Tracking Data - Enrollment data came from two sources: EnergyHub, the program implementer, and from APS directly through a system called PowerClerk. EnergyHub accepted customers into the pilot, but the interconnection process and granting of PTO went through APS. These systems required reconciliation to ensure that the customer counts aligned, unenrolled customers were captured in both system, and typos in the mapping did not prevent cross-walking between the files.
2. Delays in Telemetry Data Provision - Depending on the partner, data was provided every two weeks or once a month and it often took additional time to receive complete data for

all (or even the majority of) systems limiting the ability to conduct timely event impact estimation. Evaluators ultimately received usable data for approximately 70% of participating systems.

3. Inconsistent Field Definitions - Despite receiving a standardized data dictionary, participating battery partners provided data with inconsistent field definitions, including the sign for battery charge vs discharge and the definition of site power demand. A review of battery round trip efficiency also revealed that one partner was likely providing data on battery output power before DC to AC conversion.
4. Telemetry Data Quality Issues - Telemetry data from each battery partner suffered from idiosyncratic quality issues including systematic missing records, time zone inconsistencies, underrepresented solar production, and extreme values. Thirteen percent of systems for which data was received were removed from analysis due to data quality issues.

Figure 5 shows the percentage of systems by partner that were in the EnergyHub enrollment data (indicating systems accepted into the program), the PowerClerk enrollment data (indicating systems with PTO), the raw telemetry data, and our final analysis dataset. We show both the DO and D+M participant counts as they are both used in our savings estimation methodology. To preserve partner anonymity, we only show counts on the total section. Data cleaning performed on the delivered telemetry data to get the final analysis pool removed systems and observations for data quality issues such as missing or zero usage, negative household consumption and extreme values. Most of the full system removals were excluded because of frequent negative household consumption values. For the systems remaining in the analysis, we retained 85% of the observations after data cleaning.

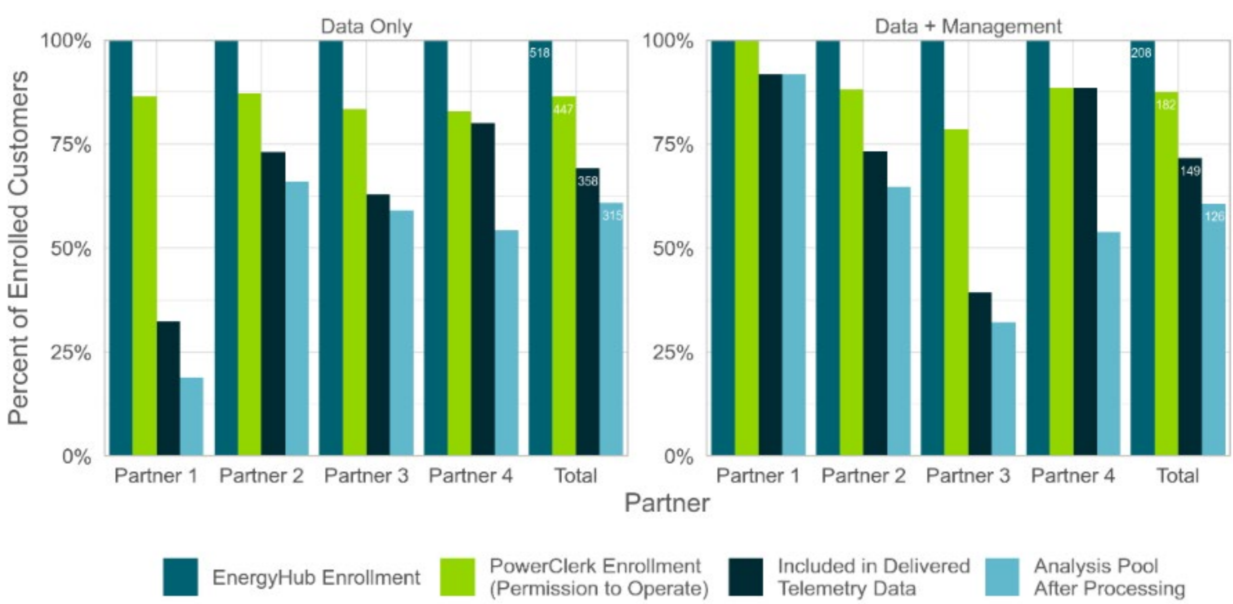

Figure 5. Data cleaning summary. *Source*: Guidehouse analysis.

Event Savings Estimation Methodology

As participants are charging and discharging their batteries daily to save money around their on-peak hours, Guidehouse estimated event savings incrementally to this daily discharge.[10] The methodology is summarized in Figure 6 below.

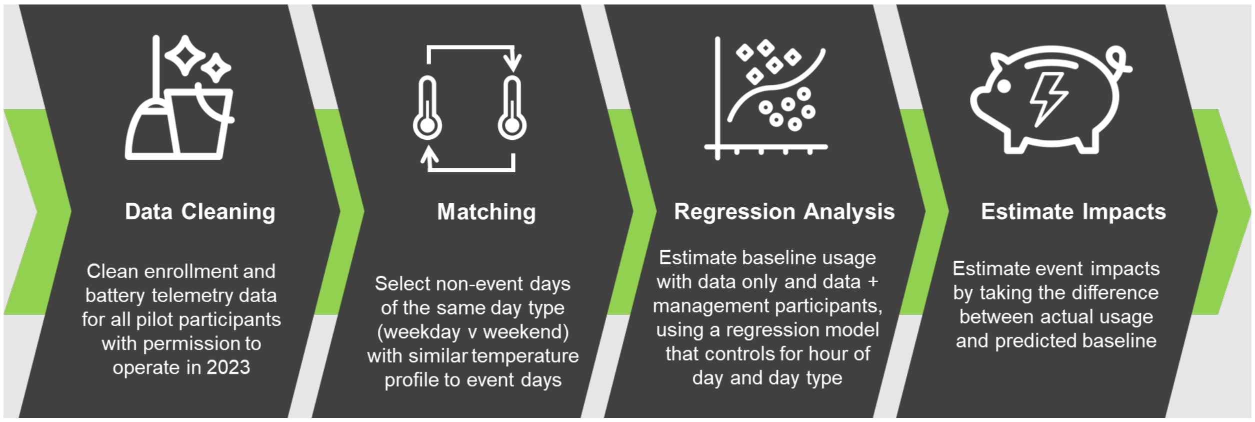

Figure 6. Event savings methodology summary. *Source*: Guidehouse.

After cleaning the data as described in the previous section, Guidehouse estimated event day impacts using a regression model where event-like, non-event days were used to construct the baseline usage. Data for both DO and D+M participants was included in the regression, but only the D+M participants were marked as having an event. Selecting "like" non-event days to build the baseline inherently captures the battery performance on event-like, non-event days and how batteries respond to parameters such as TOU rates; in this way the estimated pilot event impacts are incremental to how pilot participants use their batteries in the absence of the events.

Guidehouse selected the non-event data based on temperature profiles by day-type (weekday and weekend) to help capture usage of the home on event-like days (e.g., participants have higher usage on hotter days as their AC is running more and thus their battery discharges more to cover their home's usage during on-peak hours). Note that outside of events, the batteries do not export energy back to the grid (i.e., they can only discharge to cover the home's usage), but during events they can do so.

Guidehouse also included all DO participants in the regression to capture any differences in usage between the event-like, non-event days and the event days themselves. We did not have enough DO participants to conduct matching to those most like the D+M participants, but review of demographic breakdowns, system size, and non-event load profiles (of site usage and battery power) showed that participants across the two parts of the pilot are quite similar.

To estimate the impacts during each event, Guidehouse applied a linear regression framework using the event and selected non-event days. The regression model, shown in Equation 1, was fit separately for each battery partner after aggregating the 15-minute telemetry

[10] Note Guidehouse did not receive individual data on how participants have their batteries programmed for non-event days. However, we were told that, in general, each battery partner has similar pre-programmed setting for their customers to choose when the battery system discharges including: (1) during on-peak hours, (2) anytime the home requires more power than its solar system is producing to minimize net usage, or (3) only during outages. One partner had one battery type that was always programmed to discharge on weekdays from 3-8pm.

data to hourly. The set of coefficients β_{2jd} capture the average impact of each hour of an event, separately for each event day, by partner.

Equation 1. Regression to estimate event impacts

$$Usage_{it} = \sum_{j=1}^{24}\sum_{d=1}^{D} \beta_{1jd} Hour_{jt} \cdot EventDay_d + \sum_{j=1}^{24}\sum_{d=1}^{D} \beta_{2jd} Hour_{jt} \cdot EventDay_d \cdot Treatment_i + \beta_{3t} NonEventUsage_{it} + \varepsilon_{it}$$

Where:

$Usage_{it}$ is event day battery power consumption by system *i* in hour of the day *t*
$Hour_{jt}$ is a set of binary variables equal to 1 in hour of the day when *j*=*t* and 0 otherwise
$EventDay_d$ is a set of event-level factor variables equal to 1 during a specific event day *d* and 0 otherwise
$Treatment_i$ is a binary variable equal to 1 if system *i* is a D+M participant and 0 otherwise
$NonEventUsage_{it}$ is system *i*'s average battery power in hour of the day *t* of matched non-event days
ε_{it} is the error term, clustered at the system level

To inform whether additional battery capacity is still available following events, Guidehouse reviewed the battery availability, based on SOC, at the end of each event to see if systems are discharging down to the 20% reserve capacity prescribed by the pilot. We further assessed performance by reviewing whether systems were discharging power (kW) at the rate of their theoretical limit. The system power discharge (kW) during events should be limited by (1) the system's stored energy (kWh)/SOC and reserve capacity (equal to 20% of the system's energy storage capacity), and (2) the system's power discharge capacity (kW). The theoretical limit on system power discharge (kW) is the minimum of these two limiting factors at any point in time. This comparison of theoretical power discharge (kW) to actual looked at total discharge during the event (including both what the system would have discharged on a non-event day and the incremental event discharge).

Results

Figure 7 shows the battery power profile for two sample events, one weekday and one weekend, to illustrate how the regression and savings estimation work. The y-axis shows battery system power consumption, where positive numbers are battery charging and negative numbers are discharging. The green line indicates the baseline usage of the system in the absence of the event, as predicted by the regression. We can see this line shows discharging in the afternoon even in the absence of an event; the discharge is deeper on weekdays when the TOU peak period occurs (shaded in grey) but some discharge occurs in the baseline even on weekends. The blue line plots the actual usage of the system during the event; comparing the blue and green lines we see a similar amount of charging in the morning from solar and then deeper discharge for the event compared to the baseline during the event hours (shaded in blue). The difference between these two lines during the event hours is the predicted impact from the event and is shown by the blue bars in the figure.

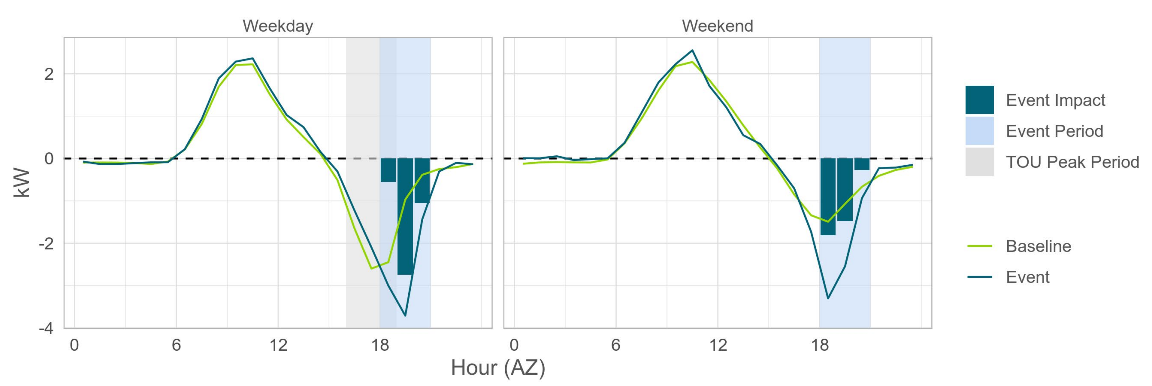

Figure 7. DR event battery profile examples. *Source*: Guidehouse.

Overall event savings are summarized in Table 2 and Table 3. Table 2 shows the average impact of an event, by event hour, averaged across all events and partners. Note that negative numbers indicate a reduction in usage and thus savings from the pilot. Impacts are highest in hour one and trail off throughout the event. Table 3 shows the highest-savings hour, out of all events, and the highest-savings hour out of all hours that fell during APS's coincident peak, which is non-holiday weekdays in July and August from 5-7 p.m..

Table 2. Average impacts by event hour

Impact hour		Average impact per system (kW)	
1	Average of 3 hours	-2.11	-1.44
2		-1.36	
3		-0.27	

Source: Guidehouse analysis.

Table 3. Highest hour event impacts

Period	Per system (kW)	Pilot Aggregate (kW)*
Highest single hour	-2.97	-540
	Sunday, September 17, 4-5 p.m.	
Highest coincident peak hour	-1.87	-342
	Tuesday, August 29, 6-7 p.m.	

* Pilot aggregate impacts equal observed per system impacts multiplied by total enrolled systems with PTO in September 2023 (182). *Source*: Guidehouse analysis.

Table 4 shows average and maximum hour impacts by partner. It is clear from these results that the various partners are performing differently. This is broken out further in the rest of the results in this section.

Table 4. Highest hour event impacts

Impact hour		Average impact per system (kW)							
		Partner 1		Partner 2		Partner 3		Partner 4	
1	Average of 3 hours	-3.68	-2.30	-1.95	-1.31	-2.16	-1.21	-2.01	-1.80
2		-2.27		-1.20		-1.19		-1.83	
3		0.38		-0.25		0.56		-1.35	
Highest single hour		-5.08 Aug. 23, 7-8 p.m.		-3.26 Sep. 17, 4-5 p.m.		-4.97 Sep. 23, 5-6 p.m.		-5.27 Sep. 22, 6-7 p.m.	

Source: Guidehouse analysis.

Figure 8 shows the average per system impact for each event hour in kW averaged by partner and month. Partner 4 underperformed particularly in July and early August, which is why we see impacts increase in September. Partner 3 also underperformed particularly in early July before picking up in later July and August. Except for Partner 4 in July, all partners are running events such that the highest impacts occur in the first hour and then trail off.

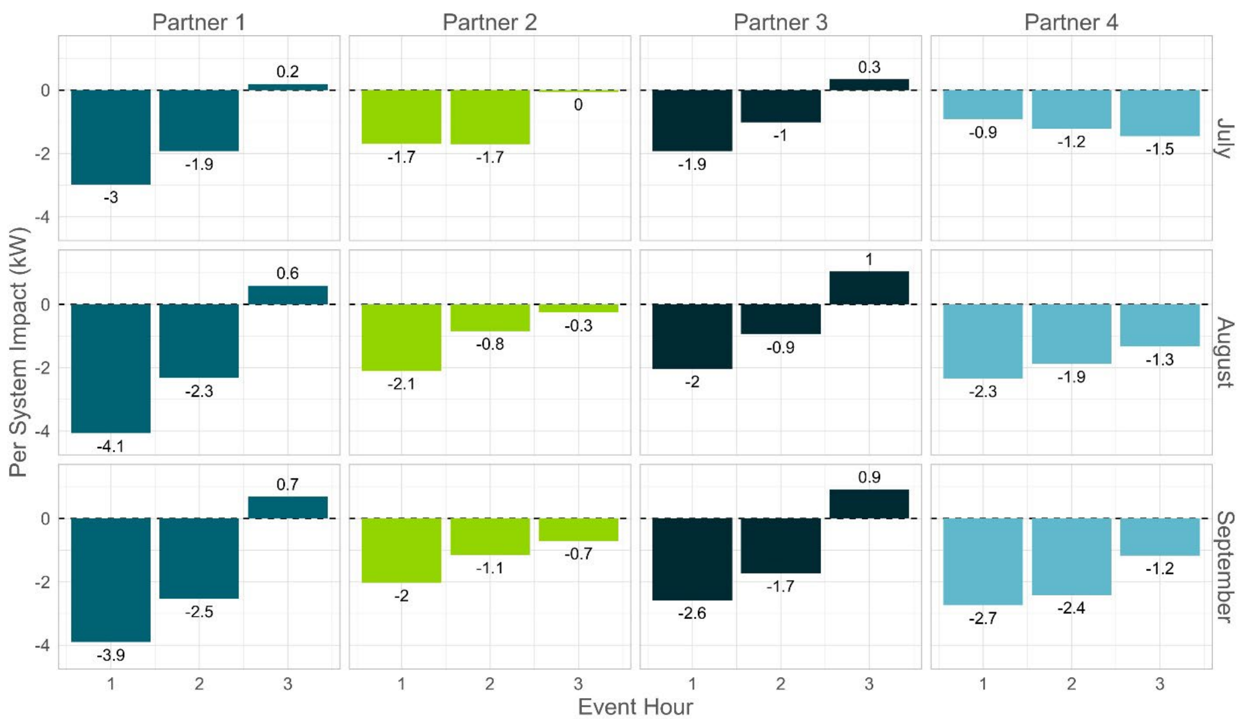

Figure 8. Average per system impacts (kW) by partner, month, and event hour. *Source:* Guidehouse analysis.

Three of the four partners had at least one event that, based on our results, failed to dispatch. An example failure date for each of these three partners is shown in Figure 9. Partner 3 failed to dispatch for six of the 29 events, a 21% failure rate, while Partner 1 failed to dispatch for three events (10%) and Partner 4 failed to dispatch for 2 events (seven percent.

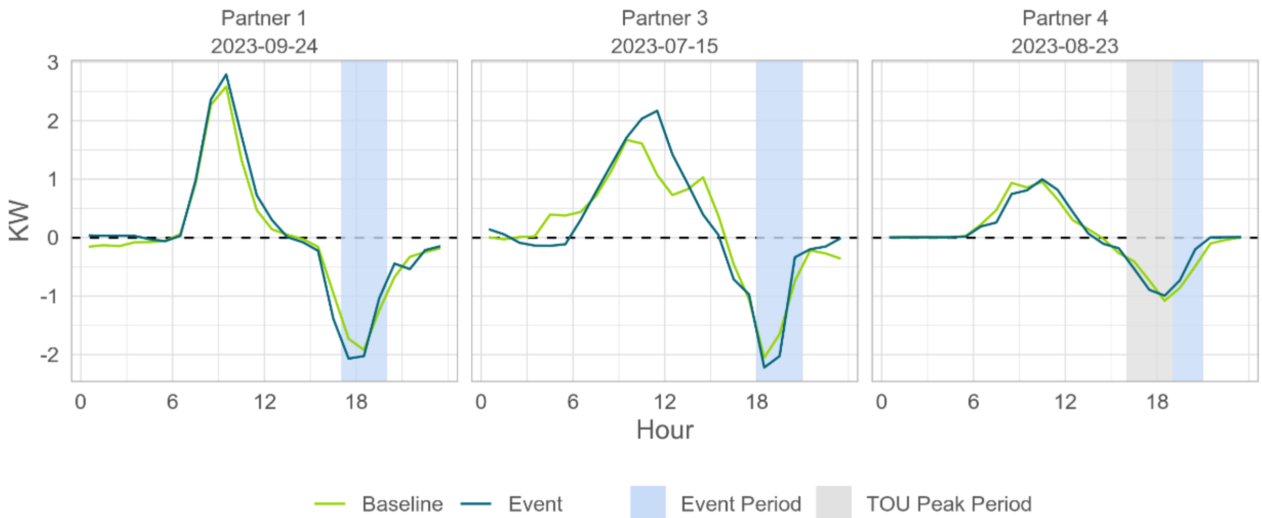

Figure 9. Battery dispatch profiles on selected failed event dates. *Source*: Guidehouse.

Figure 10 shows our analysis of the system SOC throughout an event day. To average across all events (which happened at different hours of the day), this figure is plotted by relative event hour where the event hours are shared in blue; the area to the left of the blue shading are hours leading up to the event, while the area to the right are hours following the event. The event SOC is plotted in the blue solid line while the baseline is shown by the dotted green line.

Partner 2 and Partner 3 both follow the baseline SOC quite closely up until the start of the event, whereas Partner 1 and Partner 4 diverge from the baseline in the hours leading up to the event indicating that they may be managing the systems during the pre-event hours to maintain a higher SOC up to the start of the event. This may not be desirable from a customer perspective as, for weekday events, this means the system may not be discharging during the customer's on-peak hours. This operation is counter to the APS program parameters communicated to customers when an event is called.

Additionally, the pilot reserve capacity is 20% SOC by default (though we understand that customers can override this default and set a different reserve capacity if they wish),[11] but we can see from Figure 10 that the average end-of-event SOC for systems controlled by most of the partners is above this threshold. Partner 1 does get down to 20% SOC, but Partner 2 and Partner 3 bottom out closer to 30%, and Partner 4 only gets down to about 50%. The high SOC for Partner 4 is impacted by a couple things: (1) underperformance in July and early August when SOC only went down to about 75% for many systems, and (2) even later in the season some systems only got down to 65% SOC on average; other systems routinely got down to about 30% SOC following events.

We also looked at the next day's SOC (i.e., SOC on the day after an event) to see how the recovery charging performed which typically occurs during the next day's solar period. On average, the system's recovered to their normal SOC within a few hours of solar production. This suggests that there is no impact on event performance from calling events multiple days in a

[11] The reserve capacity indicates the SOC the battery should not go below during regular operation (both within and outside of events). The remaining SOC is kept in reserve in case of an outage.

row. This also indicates that the customer's ability to discharge their battery during the next day's TOU period (outside of an event) is not impacted.

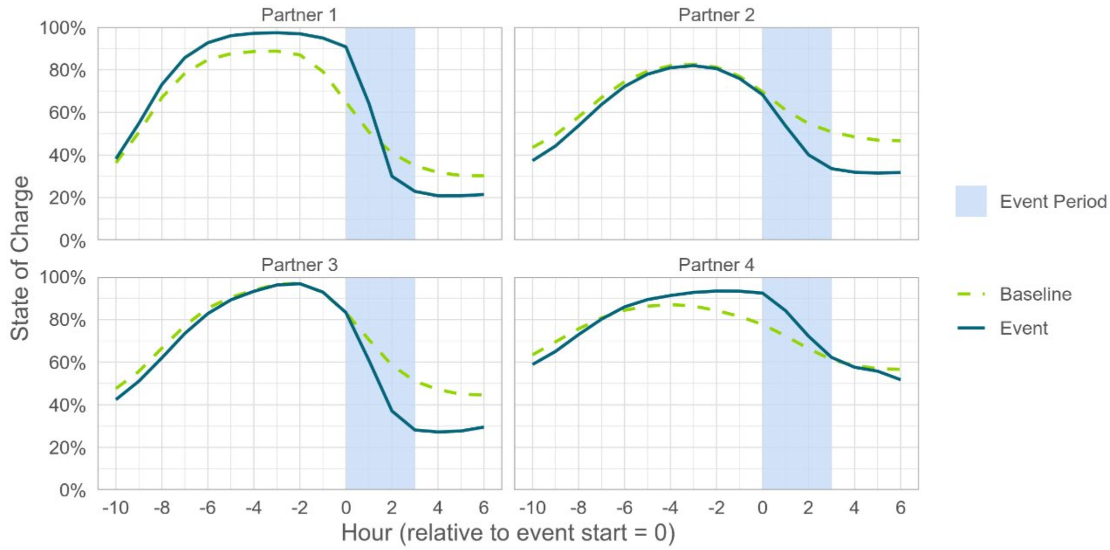

Figure 10. Battery SOC (average of all events). *Source*: Guidehouse.

Guidehouse also reviewed what an average system's power discharge (kW) was during events compared to the theoretical limit. As noted in the methodology section, system power discharge (kW) during events is limited by (1) the system's stored energy (kwh)/SOC and reserve capacity[12] and (2) the system's power discharge capacity (kW). The theoretical limit on system power discharge (kW) is the minimum of these two limiting factors at any point in time. Note that during events the power discharge (kW) is not limited to household consumption as systems can export to the grid at these times. This comparison of theoretical power discharge (kW) to actual looked at total discharge during the event (including both what the system would have discharged on a non-event day and the incremental event discharge).

We found that, during events, systems are discharging at less than 95% of their theoretical limit 49% of the time. This review also showed us that systems are charging in 36% of the 15-minute intervals during event hours. Though the charging is minimal (0.14 kW on average), it is a lost opportunity for discharge particularly for those partners who also have a SOC above 20% at the end of an event. Figure 11 shows the percent of 15-minute intervals of an event during which systems are charging, out of power,[13] and discharging compared to their theoretical limit.

[12] For the purposes of this analysis, we assumed this to be 20% per system. We were informed participants can change this reserve capacity if they want to but did not receive any statistics on how many had done so.
[13] Out of power means we do not expect a system to discharge any further because it has reached its 20% reserve margin.

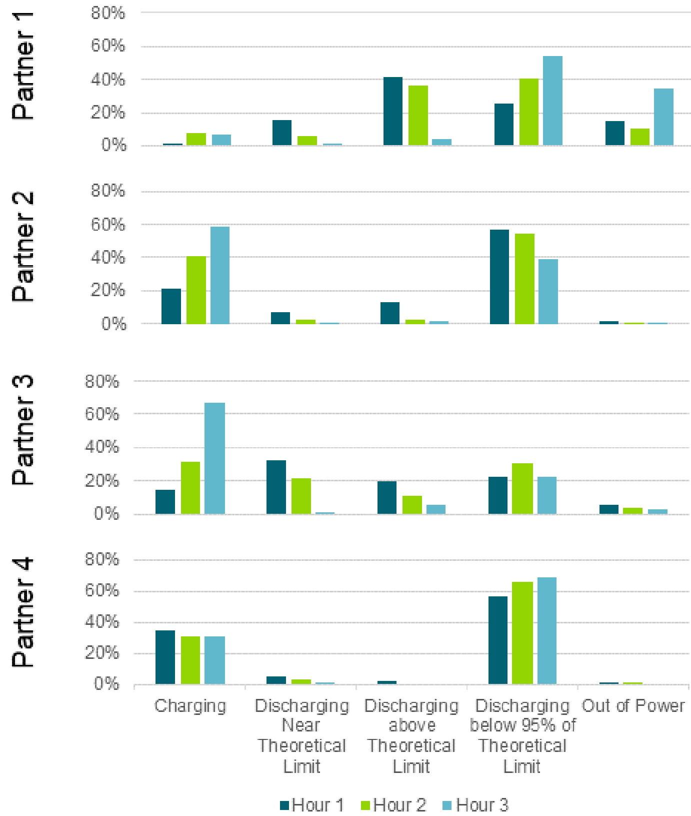

Figure 11. System behavior by event hour (average of all events). *Source*: Guidehouse.

Table 5 summarizes the percentage of 15-minute intervals of an event during which systems are discharging below 95% of the theoretical limit or charging, by partner.

Table 5. System behavior summary during event hours (average of all events)

Partner	Time spent discharging below 95% of theoretical limit	Time spent charging
Partner 1	40%	5%
Partner 2	50%	40%
Partner 3	25%	38%
Partner 4	64%	32%

Source: Guidehouse analysis.

We also broke down when systems were discharging at or below their theoretical limit into 10 buckets between 0% and 100% of the theoretical limit as shown in Figure 12. When discharging at or below their theoretical limit, Partner 3 systems are most likely to discharge at their limit rather than below it. The other partners are more distributed on discharge levels between 0-100% of their theoretical limit. For Partner 4, performance was quite variable across systems; almost 30% of discharge was between 0-10% of the theoretical limit driven by a specific subset of systems that underperformed throughout the season. Partner 1 and Partner 2 were more evenly distributed though Partner 1 had higher portions at either end of the spectrum.

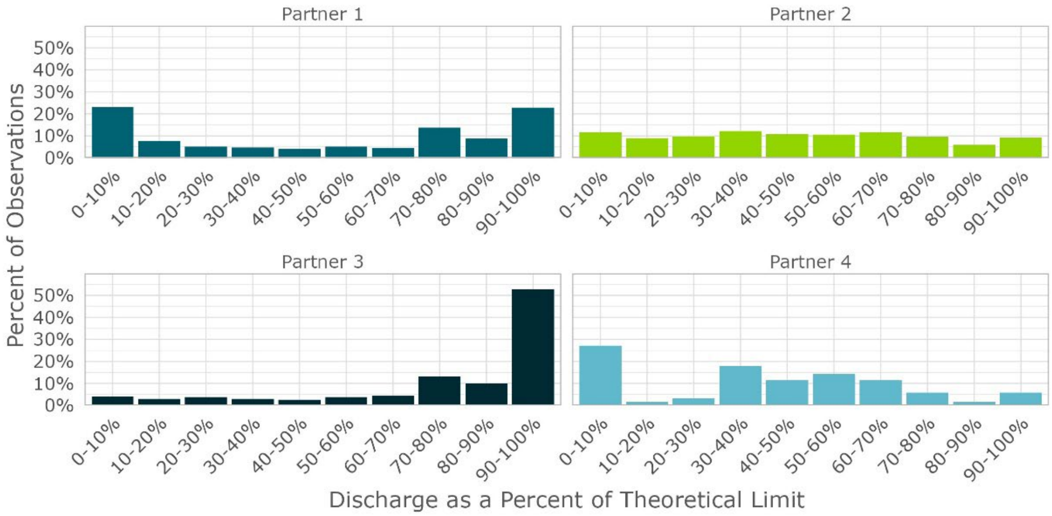

Figure 12. Discharge compared to theoretical limit (average of all events). *Source*: Guidehouse.

Conclusions

Key Findings

Our results found that battery storage systems do have charge left after an on-peak period to continue discharging during DR events. On average, weekday events, which overlap with or follow APS's on-peak period, had 1.36 kW of additional hourly discharge compared to baseline discharge behavior (i.e., how customers are otherwise discharging their battery systems on non-event days). Running these events at the end of the on-peak period allows APS to use these batteries to mitigate snapback on the grid from customers' usage ramping up when rates go down, particularly as the pilot scales. APS also ran weekend events, as they experience capacity constraints on the weekends when on-peak rates are not in effect, and saw average event impacts of 1.55 kW. Higher savings make sense on the weekend since fewer participants may be discharging their systems already in the absence of the TOU rates. These systems were able to perform for DR events even in the hottest months when participants use more of their own battery capacity to run their AC before events begin late in the afternoon – monthly average event impacts were 1.17 kW in July, 1.56 kW in August,[14] and 1.63 kW in September.

Working with multiple battery partners, even with a single aggregator, led to some challenges. Although all telemetry data was provided with the same fields, the sign conventions and definitions of those fields were not the same across all the partners leading to significant data processing to standardize the data into one cohesive dataset. Additionally, the partners had variable event results. Some partners had challenges with event failures where no impacts were apparent for specific days. Some are managing the system discharge in the hours leading up to the events which may not be desirable for participants who then aren't able to dispatch their

[14] Some of the difference between July and August is explained by certain partners who had poor performance in July as noted in the results section.

batteries during on-peak hours. Not all of the partners are discharging down to the 20% SOC reserve capacity meaning they may be able to get higher impacts if they dispatched more during events, and all of the partners (to varying extents) seem to be discharging less than their theoretical limit. Table 6 summarizes key challenges by partner.

Table 6. Challenge summary by partner

Partner	Failed Dispatch (% of Events)	Depth of Discharge (Average SOC Following Event)	Pre-Event Management	Discharge Limits	Impact Profile
Partner 1	10%	20%	Pushing additional charge on event days (closer to 100% SOC rather than 90% on non-event days) and no discharge until event	Moderate discharge below theoretical limit, about 10% of observations in 0-10% of limit	Highest in first hour, APS would prefer flexibility to shape impact as needed (e.g., flat impact through event, highest first hour, highest last hour)
Partner 2	0%	35%	None	Lots of discharge below theoretical limit; additionally, little to no discharge in the 3rd hour when there's still SOC left above limit	
Partner 3	21%	25%	None	Best performance, discharging at 90-100% of limit 50% of the time	
Partner 4	7%	50%	No discharge until event	A lot of discharge below theoretical limit	

Source: Guidehouse analysis.

Looking Ahead

APS is working to coordinate conversations with all battery partners to discuss and better understand their individual battery performance. Particularly around things like dispatch failure rates, remaining SOC after events, and optimizing discharge compared to the system's theoretical limit.

APS is considering the following recommendations for the pilot:

- **Disable any charging during events.** Even if the magnitude of charging is small, any charging at all is working against the purpose of reducing power pulled from the grid during the DR events. Additionally, most participants are better served, economically, by recharging their batteries during the next day's solar production period rather than immediately following the event. Finally, some partners seem to be charging during the events when they are still above the 20% SOC reserve capacity, indicating that these systems could be discharging more rather than charging.

- **Impacts of pre-event management**. Although managing the battery to maintain a high SOC in the hours leading up to the event can boost event performance by allowing for more discharge during the event, it also means that these batteries are not discharging in the hours leading up to the event. Particularly for weekday events, this could have an impact on customer bills since their system is not discharging during their on-peak periods leading to more usage during those times. For these reasons pre-event management is better suited as a consideration for a pay-for-performance battery management program.

Future research on the pilot will include reviewing the impacts of summer 2024 events when we should have a full summer of events, rather than starting in July, where all partners have some experience running events in this pilot, as opposed to July acting as a ramp up period where some partners saw lower impacts. Looking at impacts over two summers will allow us to confirm the consistency of the event results during the summer.

References

APS. 2023. "APS Powers Seven Days in Row of Highest Customer Electricity Use Ever." Accessed February 2024. https://www.aps.com/en/About/Our-Company/Newsroom/Articles/APS-Powers-Seven-Days-In-Row-Of-Highest-Customer-Electricity-Use-Ever

Shovelin, I. and A. Kornelis. 2023. *Battery Arbitrage: Do Customers Optimize Battery Dispatch in the Presence of TOU Rates?* Charlotte, NC: PLMA.