H4 Research Report

Incentivising within-day shifting of household electricity use

Final report

Australian Government
Department of Industry, Science and Resources
AusIndustry
Cooperative Research Centres Program

RACE for Homes

Research Theme H4: Rewarding flexicle demand

ISBN: 978-1-922746-45-0

Industry Report

Incentivising within-day shifting of household electricity use

September 2023

Project team

Monash University

- Anke Leroux
- Gordon Leslie
- Michael Ward

Citations

Leroux, A.D., Leslie, G.W., Ward, M.B. (2023). Incentivising within-day shifting of household electricity use. Prepared for RACE for 2030 CRC.

Project Partners

Acknowledgements

The Project team would like to acknowledge and thank the IRG members who provided their time and expertise during the scoping and development of the report including the review of findings and opportunities: The industry reference group comprised of individuals from Endeavour Energy, AusNet, Jemena, Network of Illawarra Consumers of Energy, DEECA (Victoria), Ausgrid, and NSW Office of Energy & Climate Change.

Acknowledgement of Country

The authors of this report would like to respectfully acknowledge the Traditional Owners of the ancestral lands throughout Australia and their connection to land, sea and community. We recognise their continuing connection to the land, waters and culture and pay our respects to them, their cultures and to their Elders past, present, and emerging.

What is RACE for 2030?

RACE for 2030 CRC is a 10-year cooperative research centre with AUD350 million of resources to fund research towards a reliable, affordable, and clean energy future. https://www.racefor2030.com.au

Disclaimer

Executive Summary

Australia's boom in rooftop solar is creating value opportunities for households to shift electricity consumption to the sunny and increasingly low-cost daytime and away from times when the sun has set. Mobilising households to undertake within-day energy use shifting is crucial for maximising the environmental benefits and economic value from renewable energy investment in Australia. Can so-called "solar sponge" tariffs or "peak shave" tariffs encourage households to meaningfully shift the timing of their energy use? If so, which tariff structures are most effective? Can it be demonstrated to electricity providers that such tariffs are worthwhile?

The objective of this project was to design and test monetary and non-monetary incentives to encourage households to shift their within-day electricity use to align with solar energy output. It involved the design, implementation and evaluation of the *Load-shifting Challenge*, a randomised controlled trial (RCT) specifically designed to test the impact of different incentive schemes on household load shifting, energy procurement costs and net program costs. The incentive schemes differed across three dimensions:

i. incentives to move consumption into daylight hours (solar sponge incentives) versus incentives to move away from non-daylight hours (peak shave incentives),
ii. financial subsidies and rebates per kWh shifted versus a non-monetary incentive scheme where households earn status points per kWh shifted, and
iii. incentives that target routine (every day) actions versus those that target ad hoc (less frequent) actions. This resulted in eight different incentive schemes that were randomly allocated to participants in the Load-shifting Challenge.

The experimental design allowed for insight into features critical to engaging households to alter their within-day electricity use activities and the economic value that it can create.

First, it directly compared the behaviour of those receiving every day versus less frequent incentives. Every day incentives, where customers are encouraged to concentrate more usage into low-cost times of day, may create value through lower average costs to service these customers. Ad hoc incentives can target high value days and provide demand response, where events (and thus usage) are flexible and driven by prevailing wholesale costs in the market. Understanding the responsiveness and habit-formation characteristics of households to routine versus ad hoc events is a crucial input to developing business cases and assessing the economic value that can be created by implementing both *demand response* and *time-of-use* programs and tariffs.

Second, the experimental design provided insights into the sensitivity of load shifting responses to incentive design. Do households respond more to solar sponge subsidies that incentivise increased consumption in the middle of the day? Or to peak shave rebates that encourage reduced consumption in the evening? And if households respond, to what extent do we observe spill-over benefits from load shifting over the daily consumption profile? For example, does evening electricity use decrease for households that receive middle-of-day solar sponge incentives?

Another important aspect of the experimental design was related to the size of the incentives offered. Is it possible to achieve load shifting by offering non-monetary incentives? Or is load shifting financially motivated? If the latter, are incentives that are in line with average price differentials between middle of the day and evening wholesale prices sufficiently high to encourage load shifting? If higher incentives are needed for households to engage in load shifting, then routine incentives may not create enough value to be commercially viable.

Lastly, given the diversity of electricity consumers it was important to understand how responses vary with household characteristics to effectively target any new program or tariff. The trial participants are equally split in terms of the number of homes with and without roof top solar installations and consequently we were able to provide novel insights into how these two customer groups respond to the same incentive programs offered. Further insights into the building form, that is free standing house through to multistorey apartments was not possible from the available data set.

Methodology

The Load-shifting Challenge was implemented as a randomised controlled trial in order to test how households respond to incentives that encourage more daytime electricity use (when the costs of generating energy are generally lower) and less evening electricity use (when the costs of generating energy are generally significantly higher). The trial participants were Powerpal users in Victoria, Australia. Powerpal is a technology provider whose customers receive a) real-time visualisations of their electricity use via a smartphone app when within Bluetooth range of their smart meter, and b) weekly energy usage reports and tips. During the trial, Powerpal users received notifications incentivising them to alter their energy use and enabling them to track their performance relative to some target usage during and after these events.

The 6,005 participants in the trial were randomly assigned into either a control group or one of eight treatment groups. The eight different incentive structures are defined by various combinations of 3 characteristics comprised of 1) whether rewards are earned by using more in the daytime (12 pm-3 pm) or less in the evening (5 pm-8 pm), 2) the size of the reward (non-monetary, 5c/kWh, 10c/kWh, 50c/kWh), and 3) the frequency of the events (every day; approximately every 3 days; approximately every 2 weeks).

Treatment impact was estimated by comparing load profiles (the average hourly electricity use pattern) between each treatment group and the control during the trial window. Statistical tests showed that the load profiles of all groups prior to the trial were indistinguishable from each other, thereby confirming successful randomisation. Analyses on the effect of household characteristics examined whether the responses to the treatments differed between households with and without solar panels, and leveraged self-reported survey data on load-shifting techniques and trial experiences to provide insights into whether these factors varied across household characteristics and treatment allocation.

Findings

Routine vs ad hoc programs

The results of the Load-shifting Challenge demonstrate that simple and small incentives can change aggregate consumption profiles among trial participants independent of being offered every day or on an ad hoc basis. As ad hoc incentives changed consumption profiles on event days, this demonstrates the potential for these incentives to generate flexible demand response.

The observed responses to routine incentives suggest that time-of-use-style tariffs can change load shapes and lower the average cost of supplying households. This demonstrates the potential for a retailer-customer surplus if customers move from time-invariant fixed rates to time-of-use fixed rates. However, these tariffs do not provide demand response because they do not promote flexibility with respect to real-time conditions. In contrast, ad hoc events can be called when market or network conditions are such that changing load profiles is forecast to be particularly valuable on specific days. These results demonstrate meaningful consumption responses can be generated by households on ad hoc event days, and of similar size to responses generated by households that receive every day incentives.

Solar and non-solar households

Non-solar households were more sensitive to incentive design than solar households, changing their pattern of electricity use in response to some, but not all program designs. In contrast solar households responded to all program designs.

For example, non-solar households did not respond to every day monetary or non-monetary peak shave incentives. Solar households, in contrast, responded to either form of incentive (decreasing their peak usage by 9%), demonstrating no obvious sensitivity between being paid nothing or 5c/kWh to shave their peak consumption.

Non-solar households did, however, respond to monetary incentives where they were paid to use more energy in the middle of the day (solar sponge incentives). Specifically, they increased their energy during the middle-of-day incentive window by an average of 6% and interestingly, also decreased energy consumption during the evening peak by an average of 8% despite not receiving a direct incentive to do so. This change in the consumption profile is shown in Figure E-1, where the black line is the control group's average consumption profile and the red line the average consumption profile of households receiving the solar sponge incentive. Notice the change in trajectory of usage at midday, and also the reduction in energy use in the evening. This increase in low-cost middle-of-the-day electricity use and the accompanying decrease in high-cost evening use is an encouraging result for policy-makers and retailers that aim to reduce the average cost of energy for these customers.

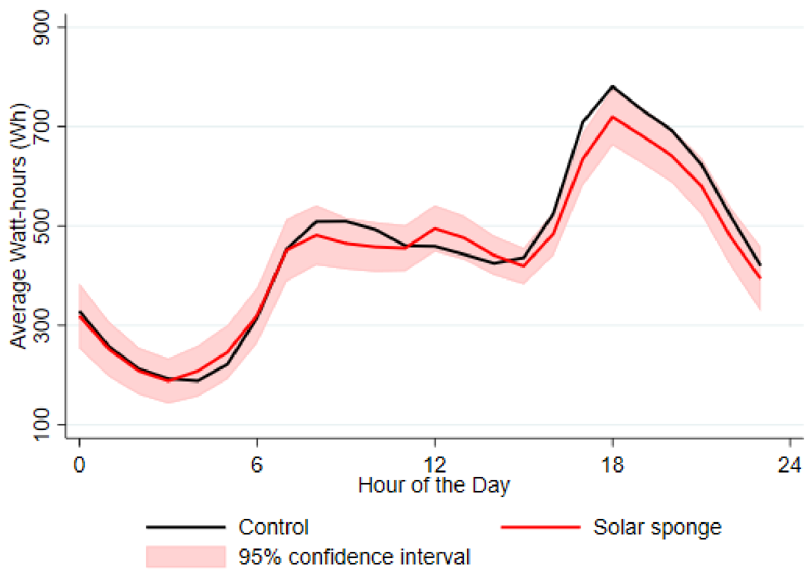


Figure E-1: Average treatment effects for routine, solar sponge, 5c/kWh incentives: Non-solar users; details for figure are described in Sections 7 and 8

Finally, non-solar households were found to respond to peak shave incentives only when the reward was substantial (50c/kWh). In short, non-solar households engaged with and responded to some but not all load shifting treatments, highlighting that careful incentive design and targeting can improve program benefits.

Value creation

Participant consumption profiles were impacted by many of the treatments trialled, and the results showed that under certain wholesale price conditions these treatments could generate financial benefit for both households and retailers. Under prevailing wholesale market conditions, and despite the short duration of the trial, some treatments decreased wholesale procurement costs, providing a simple demonstration of the potential for financial benefit. For example, it was estimated that wholesale procurement costs for non-solar households that received every day solar sponge incentives were $29 (9%) lower than for the control group over 90 days, or $22 (7%) lower when including the program payments to these households.

Glossary

AEMO	Australian Electricity Market Operator
CP	Critical Peak
DELWP	Department of Environment, Land, Water and Planning (Now Department of Energy Environment & Action)
DNSP	Distribution and Network Service Provider
DR	Demand Response
GW	Gigawatt
kWh	Kilowatt-hour
MW	Megawatt
MWh	Megawatt-hour
Program	Programs or trials that attempt to shift electricity consumption using incentives
Project	The Load-shifting Challenge project from funding to information dissemination stage
PV	Photovoltaic
RCT	Randomised controlled trial
ToU	Time of Use
Treatment	A specific incentive design that was tested in the trial
Treatment group	A group of trial participants who are randomly assigned to a treatment
Trial	Time period when Load-shifting Challenge participants were actively treated
W	Watt
Wh	Watt-hour

Contents

1 Introduction

Solar PV systems are an integral part of Australia's energy transition. Since 2010, both small-scale rooftop and larger commercial and utility-scale installations have grown rapidly in number and size. Figures from March 2023 show that Australia is amongst the world leaders in solar penetration, totalling over 3.44 million PV installations country-wide with a combined capacity of over 30.5 GW (Australian PV Institute 2023, https://pv-map.apvi.org.au/analyses).

The boom in solar installations is accompanied by two wholesale price trends in the national electricity market. On the one hand prices continue to fall during daytime hours, while on the other, prices are increasing in the evening. These trends reflect falling costs of generation in the middle of the day when output from our solar resources is high, but increasing costs of generation in the evening, when PV systems cease generating and this reduction in electricity generating must be met by wind, gas-fired or other higher-cost generators to cover evening peak demand.

At the same time, the renewable energy transition leaves some parts of the electricity network more vulnerable to day-to-day instabilities. The Australian Energy Market Operator (AEMO) and distribution and network service providers (DNSPs) already view low minimum demand (that is, total electricity use less the amounts generated by roof top solar) levels in the middle of the day as an operational issue that is projected to become increasingly problematic in the near future (AEMO, 2022).

Another implication of the increased share of renewables in the energy mix is that their intermittency causes energy supply and, in the case of rooftop solar, net energy demand, to become more volatile and difficult to predict. This increases the value of solutions that can support flexible responses to ameliorate this volatility. For example, households supplying less/using more energy when renewable output is unexpectedly high can help balance network supply and demand.

The price differential between daytime and evening peak prices creates value opportunities for electricity users to shift more consumption to the sunny daytime and away from times when the sun has set. Figure 1-1 documents how average wholesale prices have changed in our setting (Victoria, Australia) from 2015-2020, aligning with a boom in new rooftop solar installations of 0.3 GW in 2015-16, 1 GW in 2017-18 and 1.9 GW in 2019-20. The difference in average midday to early-evening prices has increased from roughly $25/MWh in 2015-16 to $40 in 2017-2018 to $100 in 2019-20. Although stark, this figure does not depict the increase in volatility and understates the value to be gained by targeting days with very high price differentials compared with every day programs that also cover days when price differentials are zero or even negative. For example, in 2019-20, load shifted from 6pm to midday would have created 17c/kWh for the top 25% days in terms of price differential, 34c/kWh for the top 5% of days, and 200c/kWh for the top 1% of days.

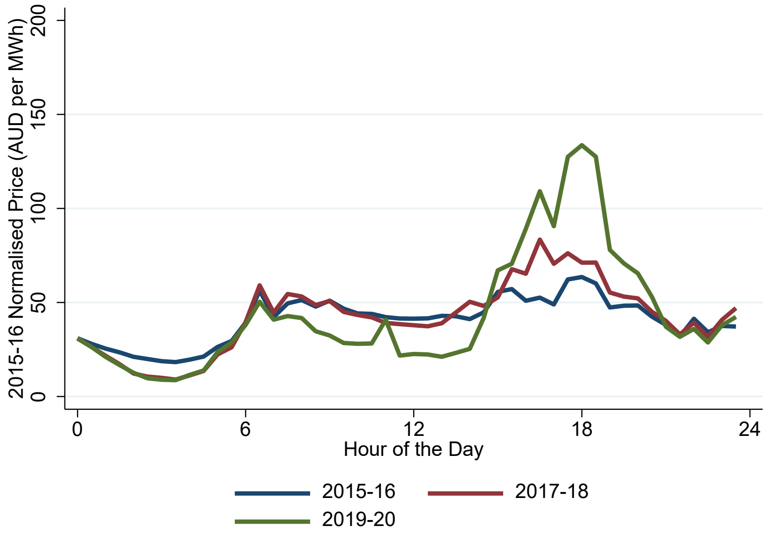


Figure 1-1: Average Victorian wholesale electricity price by half-hour of day (Normalised by hour 0, 2015-16)

At scale, within-day shifting of household electricity use leads to an increase in market demand in the middle of the day and less demand in the evening. Beyond the clear economic value that can result from such actions by lowering the costs of supply, there may also be impacts on carbon emissions. Additionally, within-day shifting may also contribute to grid stability if it alleviates voltage control issues in the distribution network derived from low levels of minimum demand and subsequently high levels of solar feed-ins.

However, there are challenges to unlocking this value, with few real-world examples to demonstrate to either retailers or households that they can benefit from offering or facing tariffs designed to encourage load shifting. What will it take to engage households with so-called solar sponge tariffs to meaningfully shift their timing of energy use, and can it be demonstrated to electricity providers that such tariffs are worthwhile?

An obvious barrier to households engaging in load shifting is the absence of financial incentives under the fixed-rate tariffs that most face.[1] Under a fixed-rate tariff, even if households are willing to shift consumption from the evening to the daytime, the value of doing so is captured solely by their retailer. Although time-of-use tariffs might create value by shifting average consumption profiles, they cannot provide true demand response to changing market conditions as they arise because consumers are not exposed to real-time price signals. Hence, it is important to distinguish between the potential barriers to engaging households on both time-of-use programs and ad hoc programs that can provide demand response.

On the flipside, energy retailers are perhaps slow to offer these programs due to insufficient evidence that customers are willing to engage with time-varying prices designed to encourage within-day consumption shifting. Further, even if customers do engage, do their resulting actions generate a sufficient financial benefit that can be shared between the retailer and customer?

In a randomised controlled trial (RCT) with over 6,000 Victorian households, this research investigated two central questions: 1) Are households able and willing to engage in within-day load shifting if incentivised to do so, and 2) Which tariff designs are the most effective?

[1] 87% of metered household loads are reported to face a tariff with no time-of-use characteristics in (CME, 2017).

2 Project Scope

2.1 Project aims

The objective of this project was to design and test monetary and non-monetary incentives to encourage households to shift their within-day electricity use to align with solar energy output. In particular, the aim was to quantify the impact that different incentive designs have on household load shifting, energy procurement costs, and net program costs. Trial incentives varied across three dimensions. The first dimension related to the action that is targeted by the incentive, with households being rewarded either for moving consumption into daylight hours (solar sponge incentives) or for moving consumption away from non-daylight hours (peak shave incentives). Secondly, incentives varied in their intensity with rebates and subsidies ranging from 5c to 50c per kWh shifted as well as a non-monetary incentive scheme whereby households earnt status points per kWh shifted. The third dimension related to the frequency of actions targeted, with some trial participants being incentivised daily to encourage routine actions while others were offered incentives on an ad hoc basis. In total eight different incentive schemes were designed and tested in a RCT comprised of eight treatment groups (treatments) and a control group.

The experimental design allowed for insight into three crucial features on the ability to engage households to alter their within-day electricity use activities and the economic value that it can create.

First, it directly compared the behaviour of those receiving every day versus less frequent incentives. Daily incentives, where customers are encouraged to concentrate more usage into low-cost times of day, may create value through lower average costs to service these customers. Ad hoc incentives can target high-value days and provide demand response where events (and thus usage) are flexible and driven by prevailing wholesale costs in the market. Understanding the responsiveness and habit-formation characteristics of households to routine versus ad hoc events is crucial to developing business cases and assessing the economic value that can be created by implementing both *demand response* and *time-of-use* programs and tariffs.

Second, the experimental design provided insights into the sensitivity of load-shifting responses to different incentives. Do households respond more to solar sponge subsidies that incentivise increased consumption in the middle of the day or to peak shave rebates that encourage reduced consumption in the evening? And if households respond, to what extent do we observe spill-over benefits from load shifting over the daily consumption profile? For example, does evening electricity use decrease for households that receive middle-of-day solar sponge incentives?

Another important aspect of incentive design relates to the size of the incentives offered: Is it possible to achieve load shifting by offering non-monetary incentives or is load shifting financially motivated? If the latter, are incentives that are in line with average price differentials between middle-of-the-day and evening wholesale prices sufficiently high to encourage load shifting? If higher incentives are needed for households to engage in load shifting, then routine programs may not create enough value to be commercially viable.

Finally, the trial participants in almost equal numbers have homes with and without rooftop solar installations and consequently we were able to provide novel insights into how these two customer groups respond to the same incentive programs offered. Examining how responses vary across key household characteristics can help effectively target any new program or tariff.

2.2 Project partners

Powerpal: Powerpal is a technology provider offering a mobile app and supporting low-cost hardware that allows real-time energy data to be collected from existing digital energy meter assets. In the State of Victoria (our research setting), all households have “smart” interval meters, and those choosing to install Powerpal’s device can do so free of charge because they are fully subsidised under the Victorian Government’s energy upgrades scheme. Powerpal’s app provides data visualisations of the households’ energy use and allows customers to wirelessly track how much energy the home is using in real-time. The company also provides weekly consumption reports and customised tips and guidance to its 100,000+ customers. Powerpal designed new app interfaces for each of the treatment groups and managed the invitation, recruitment, trial phase, exit survey, data collection and incentive payments to the trial participants through the Powerpal app. Powerpal provided hourly interval net electricity-consumption data from participating households in anonymised form.

DELWP: The Victorian Department of Environment, Land, Water and Planning participated in the Industry Reference Group and provided feedback on the proposed research and outcomes from a public policy point of view. As of May 20323, now known as the Department of Energy, Environment & Climate Action.

2.3 Project outputs

The research project resulted in a number of tangible outputs.

Output 1: This final report, containing an analysis of the trial, documenting the impacts each load-shifting incentive has on within-day shifting of electricity use and the energy cost implications for retailers, both in terms of electricity procurement costs and program costs. In particular, electricity use prior to and during the event windows are analysed, as is any evidence of habitual effects during the trials.

Output 2: Two presentations to the industry reference group followed by feedback and discussion in a workshop format. The first workshop was held at the beginning of the trials and was used to introduce the project and seek the IRG’s input into the exit survey. The second workshop was held after the end of the trials and before submitting the final draft report and served to present and receive feedback on the preliminary outcomes and findings of the load-shifting trial. The industry reference group comprised of individuals from Endeavour Energy, AusNet, Jemena, Network of Illawarra Consumers of Energy, DEECA (Victoria), Ausgrid, and NSW DPE (now Office of Energy & Climate Change).

Output 3: Work is currently underway to prepare manuscripts of this research for publication in peer reviewed economic journals and for presentation at academic conferences and workshops.

3 Related Literature

This project leveraged insights from and contributes to the literature on dynamic electricity pricing, habituation and habit formation, and demand response studies undertaken in Australia and overseas. Our experimental design is detailed in Section 4, but where appropriate, we describe in this section how the features of our study differ from key studies in the literature.

3.1 Dynamic electricity pricing

3.1.1 Evidence on price sensitivity of electricity demand

Since the early 2000s, overseas energy retailers have been experimenting with time-varying tariffs to test whether electricity consumers respond to price incentives by reducing their consumption during peak times. Time-varying pricing schemes broadly fall into three categories:

i) time-of-use (ToU) pricing, where the price of electricity changes consistently at set times of the day (for example during peak and off-peak periods),

ii) real-time pricing schemes that allow retail prices to vary with the wholesale electricity price at hourly or shorter intervals, and

iii) critical peak (CP) pricing, whereby large jumps in the price of electricity are used infrequently to incentivise reduced energy consumption during periods of critical peak demand.

To date, most time-varying pricing trials fall within the ToU and CP pricing categories. Assessing the outcomes of 15 of the earliest pricing studies in the U.S., Faruqui and Sergici (2010) reported that participants in these studies exhibited price elasticities of daily energy demand of between -0.02 and -0.10 and had on-peak/off-peak substitution elasticities ranging from 0.07 to 0.40. Greater demand responses are observed if the home has central air conditioning and/or was provided with enabling technologies.

A review of 32 European dynamic pricing schemes provided further evidence that electricity demand is price elastic in a wide variety of settings (Kessels et al., 2016). A qualitative comparison across different schemes reveals that, in the absence of automation, offering simple and predictable schemes is important in the case of ToU tariffs and timely announcements matter in the case of CP pricing. Overall, while larger price spreads were found to result in larger consumption shifts, the current evidence suggests that large price differentials are not essential to achieving consistent, albeit smaller, shifts in electricity consumption.

The results from RCTs examining demand response are roughly in line with empirical estimates of a short run price elasticity of electricity demand of around -0.1 (Allcott, 2011; Faruqui & Sergici, 2010; Jessoe & Rapson, 2014; Wolak, 2011). Interestingly, large and persistent changes in electricity prices have been shown to result in demand becoming more elastic over time, with one study estimating the price elasticity of demand at -0.27 two years after the price change, compared with an elasticity of -0.09 six months after the change (Deryugina, MacKay, & Reif, 2020).

Combining the existing evidence on the price elasticity of electricity demand with the requirement that the incentive design be commercially viable, four levels of financial incentives were compared in the Load-shifting Challenge, with marginal incentives corresponding to 0c/kWh, 5c/kWh, 10c/kWh and 50c/kWh. Meaningful load shifting was observed in response to these comparatively low levels of financial incentives offered. However, these responses were quantitatively similar across treatments, with price elasticity only weakly observable amongst non-solar customers.

The reporting of aggregate results masks important heterogeneity in the consumption response to time-varying pricing. Firstly, the responsiveness of households to varying prices depends on their characteristics: Larger consumption effects are observed for high-use households and those with smart technologies, while low-income households tend to save proportionally more on their electricity bill (Herter, 2007; Wolak, 2010). Besides leveraging ordinary household characteristics, the Load-shifting Challenge was able to provide novel insights into how the responses differ across households with and without rooftop solar installations.

Secondly, the design of the pricing scheme was found to matter. For example, comparing hourly pricing, CP pricing and critical peak rebates within a pricing experiment in the U.S. District of Columbia revealed that, for the same price increase, hourly real-time pricing results in comparable reductions in electricity consumption than during peak periods in CP pricing. In contrast, comparable peak rebates elicit a significantly smaller response (Wolak, 2007, 2010). Also, monetary incentives have been shown to work better when combined with information treatments. In an RCT designed to test the effect of the frequency of providing information on residential electricity usage, Jessoe and Rapson (2014) found that informed households were three standard deviations more responsive to temporary price increases than uninformed households.

In the Load-shifting Challenge, both solar and non-solar households responded to ad hoc and routine incentives. However, non-solar households responded more strongly to solar sponge incentives, while only solar households responded consistently to peak shave incentives.

When implemented on their own, non-monetary interventions that provide information or use moral suasion have been found to be less effective than financial incentives (Burkhardt, Gillingham, & Kopalle, 2019; Ito, Ida, & Tanaka, 2018). For example, Burkhardt et al. (2019) randomly assigned households to peak-pricing and non-monetary incentives, including information interventions and conservation appeals during peak load events. While their financially incentivised treatment groups consumed on average 14% less electricity, they found only a minimal consumption response in the non-monetary treatments.

Another example of a non-monetary intervention is the Opower Home Energy Report, whereby thousands of randomly selected households received information on the previous months' electricity consumption and how this consumption compared to that of their neighbours. Households have been found to reduce their electricity consumption by around 1%-2% after receiving the home energy report (Allcott & Rogers, 2014; Brandon, List, Metcalfe, Price, & Rundhammer, 2019). When such social nudges are combined with targeted appeals during peak load events the effect has been much higher, resulting in up to 7% reduction in electricity use (Brandon et al., 2019).

The Load-shifting Challenge provided novel insights into the effectiveness of real-time feedback provided in conjunction with monetary or non-monetary incentives. The Challenge featured real-time feedback via the Powerpal app about the household's electricity consumption relative to the household's baseline, as well as the amount of financial incentives or non-monetary stars earnt as the event progressed and cumulatively over all events to date.

3.1.2 Evidence on load shifting

The increased uptake in renewables has had profound impacts on the daily profile of net market demand for electricity. The duck curve is a well-documented phenomenon whereby the abundance of solar energy creates a net demand trough and results in low electricity prices during the middle of the day. The trough is followed by a steep ramp up in electricity demand from the grid that is reflected in higher prices in the evening, when intermittent renewable sources of energy are replaced by dispatchable (often gas-fired) generators (Bushnell

& Novan, 2021; Jha & Leslie, 2021). This development reaffirms the benefits that dynamic pricing incentives might unlock, not only to reduce peak demand but also for alleviating network problems associated with low demand. While early studies typically do not observe increased demand outside of the peak price periods (Allcott, 2011; Wolak, 2007), more recent studies have added off-peak incentives designed to stimulate off-peak demand. For example, offering electricity at 25% of the normal rate at night, when wind generation is high, results in similar responses in absolute terms compared to raising the price five to six times the normal rate during critical peaks in one context (Burkhardt et al., 2019). In a study that focused specifically on the comparison of incentives provided to shift electricity into event windows with those that incentivise shifting consumption away from event windows, Andersen, Hansen, Jensen, and Wolak (2021), observed consumption shifts that were 2-3x greater for the *Into* treatments as compared with the *Away* treatments. In addition, *Into* treatments resulted in reduced demand before and after the event window, suggesting that incentives to increase consumption during periods of low demand may also benefit the grid during periods of high demand. Andersen et al. (2021) also found that comparatively small incentives of 5%, 20% and 50% rebates are sufficient to create value for consumers and retailers in Denmark.

The Load-shifting Challenge also tested the effectiveness of providing incentives to increase consumption versus incentives to reduce consumption during certain times of the day. Unlike Andersen et al. (2021), the Challenge focused on two specific time windows: a solar sponge window in the middle of the day and a peak shave window in the early evening. This design enabled more precise estimates of the effects of solar sponge and peak shave incentives on demand during the time window that is not directly targeted. In particular, it was observed that incentivising more electricity use during the middle of the day resulted in reduced use in the evening, while incentivising reduced energy use in the evening did not result in increased use during the day. In addition, the Challenge differs from Andersen et al. (2021) in that load shifts are compared to a true control group consisting of Powerpal customers who signed up for the Challenge but were not accepted to take part in the trial and never received any guidance or incentives to load shift.[2] This design feature meant that we could investigate short-term habit formation in ad hoc groups by comparing their electricity use on non-event days to that of our control.

3.2 Habituation and habit formation

As interventions were repeated, an interesting question arose: Does repetition desensitise the recipient, resulting in the effect of the intervention wearing off over time (Thompson and Spencer 1966), or does it help the formation of new consumption habits that have lasting effects even after the intervention is discontinued (Becker & Murphy, 1988; Stigler & Becker, 1977)?

There is strong evidence of persistence in relation to electricity consumption. For example, Costa and Gerard (2021) studied the long-term effects of an energy conservation policy that was implemented in affected areas during a period of drastic supply shortage in Brazil. The 9-month policy introduced consumption quotas and additional financial incentives to encourage consumption below the quota, resulting in a 23% reduction in electricity consumption on average. About half of this effect remained twelve years after the policy ended.

Similarly, the energy conservation effects of the Opower Home Energy Reports exhibit persistence. Allcott and Rogers (2014) found that consumers are slow to habituate to the receipt of new reports, resulting in significant action followed a period of backsliding where recipients revert to their previous consumption

[2] In Andersen et al. (2021), the timing of load-shifting events is randomised across trial participants, with the control group for any particular event being made up of participants who were not notified of that event.

habits. While these cycles weaken over time, they do not disappear completely, even after five years of receiving the monthly reports. Once the reports are discontinued, the energy conservation effect decays at a rate of 10%-20% per year. In contrast, Ito et al. (2018) found that households desensitise quickly to repeated moral suasion interventions to reduce energy consumption during peak demand hours, but that the original treatment effect can be restored after a sufficiently long break between interventions. Repeated financial incentives in the same experiment, on the other hand, result in continued energy conservation even after the incentives are discontinued.

Persistent conservation has been observed in relation to household energy as well as water consumption, raising the question whether habit formation or technology adoption is responsible for this persistence. Brandon et al. (2022) analyse the energy consumption of households that moved into homes where the previous occupants received Opower Home Energy Reports. They find that, after the change in occupants, just over half of the long-term reduction in energy use attributable to the receipt of the home energy report remains in the home, despite the new occupants not receiving the report. Based on this finding, Brandon et al. (2022) concluded that technology adoption, including fixed energy efficient investments, are the primary channel through which the Opower report achieves persistent energy conservation outcomes. A number of other studies that investigate the same question favour habit formation as being key to persistent reductions in electricity consumption (Allcott & Rogers, 2014; Costa & Gerard, 2021; Ito et al., 2018).

If these interventions do change consumption habits, what is the mechanism by which this is achieved? Building on the economic (Becker & Murphy, 1988; Stigler & Becker, 1977) and neuropsychological literature (Anderson, 2016; Anderson, Laurent, & Yantis, 2011), Byrne et al. (2021) devised an experiment to disentangle a consumption-based mechanism from an attention-based mechanism. The consumption-based mechanism due to Becker (1977) describes habit formation as a process where changes in consumption over time are becoming ingrained as consumption habits. In contrast, the attention-based model attributes the benefit of repeated interventions to providing recurring cues to paying attention to resource use, which results in the formation of an attention-paying habit. The experiment involved a shower device that is pre-programmed to provide real-time feedback on shower water use at varying intervals over six weeks. Consistent with the predictions of an attention-habit model, they found that this feedback resulted in an immediate and stable change in behaviour that eroded gradually as the feedback was turned off.

By investigating consumption on non-event days across routine and ad hoc treatment groups in the Load-shifting Challenge, this research was able to provide novel insights on the relative effectiveness of offering real-time information on resource use (i.e. an attention-based mechanism) in conjunction with either monetary or non-monetary incentives that are repeated continuously versus in an ad hoc way in creating persistent load shifting habits.

3.3 The Australian context

Against a backdrop of high rooftop solar penetration and frequent occurrences of high evening wholesale prices, there is significant interest in reducing peak demand in the Australian market. At the same time, few Australian households have opted into retail contracts featuring time-varying prices and therefore receive no direct price signal to help them align their electricity consumption with prevailing market conditions. Retailers and network companies have responded by trialling demand-response programs that are focused on reducing peak demand by offering a variety of monetary and non-monetary incentives to households. These programs mostly target demand during critical peaks, using ad hoc day-ahead and hour-ahead messaging. The RACE for

2030 Opportunity Assessment: *Rewarding flexible demand* (Roberts et al., 2021) provides an overview of these programs that is summarised and reproduced here in Table 3-1 below.

The financial incentives offered in the Australian demand-response programs typically entail a participation payment and/or payments that are tied to the household's demand response relative to an estimated consumption baseline.[3] Some programs feature information treatments or moral suasion in conjunction with monetary incentives, while one program relies solely on non-monetary incentives. In general, trial participants report to be more attracted to financial rewards than non-monetary incentives. That said, many participants consider the value of the incentives offered in these programs to be insignificant and cite non-financial motives for their participation. Despite participants perceiving the incentives offered as low level, the value of demand response delivered often fails to cover the cost of the monetary incentives.

Most trials report consumption responses to their various incentive designs, with households reported to reduce their electricity consumption during peak demand windows by between 0.24 and 0.90kW on average, with high heterogeneity observed across households. Trials with near real-time feedback and variable incentives generally report greater achieved demand response. The trials also find tentative evidence of limited load shifting with increased consumption outside of event windows observed in some trials.

The Opportunity Assessment *Rewarding flexible demand* (Roberts et al., 2021) identifies several priority research areas that are relevant in the context of the Load-shifting Challenge. In particular, it identifies a need for more research into how households understand their own electricity consumption, and how this understanding is impacted by having access to real-time energy feedback and by their participation in demand-response programs and trials. The report also identifies the need to be able to differentiate consumption decisions for households that comprise of consumers only from those that also produce electricity (prosumers). Finally, the report also recommends investigating mechanisms that can increase midday demand and their value.

The Load-shifting Challenge comprised several novel features in the Australian context.

Firstly, participants in the Load-shifting Challenge could monitor their electricity consumption in real-time via the Powerpal app. The app's functionality was augmented for the trial to allow participants to track their consumption and reward earning to date and for each event against their personalised baseline.

Secondly, the Load-shifting Challenge incentivised consumption shifts into the middle of the day as well as away from early evening peaks, thereby addressing not only the need to manage evening peak demands but also the increasing concerns around demand troughs in the middle of the day. Importantly, the experimental design enabled the systematic comparison of solar sponge and peak shave incentives in terms of their impact on consumption during the event window, as well as any load shifting that occurs before or after the event.

Thirdly, the participants in the Load-shifting Challenge were almost evenly split into solar and non-solar households enabling unique insights into how these two customer segments respond to different incentive designs.

Finally, in addition to understanding demand response to rare critical-peak incentives, a key objective of the Load-shifting Challenge was to examine the potential to achieve changes in consumption habits that result in

[3] Payments under the latter category consist of target-based bonuses and rebates of between $1.50 to $5.00 per kWh of reduced consumption or a combination thereof.

increased demand in the middle of the day and lower evening peaks every day. To this end, different mechanisms to motivate the formation of such habits were tested and compared.

Table 3-1: Summary of demand response trials and programs in Australia. Source: Roberts et al. (2021)

Organisation	Project	Date(s)	Notification	Incentives	Baseline / Verification	Messaging	Monitoring technology	Key finding(s)
Endeavour	PeakSaver / CommunitySaver / PowerSaver	-> Current	In-app notification	$1.50/kWh gift card. Suggest using cinema tickets to incentivise leaving the house during events	Based on demand in preceding business days adjusted to use on day of event	Some users respond without financial reward, but distrust is an issue. Community challenges: "learn and earn", make a pledge", "take an action"	Smart Meter and mobile app	Not attractive to energy efficient households. High cost of marketing and education.
Energy Australia	Mass Market / Energy Saving Reward Program / Power Response	2017-> Current	SMS / Email	Now: $1/event participation + $2/kWh bill credits	AEMO adjusted 10-day average baseline. But it's not representative and not suitable for solar households	Effective household engagement is paramount and ongoing – education, branding, style, safety and voluntary nature are important"	Smart meter	Financial rewards were more appealing than altruistic rewards for MM households. Low smart meter penetration is challenging.
Enova Energy	Voluntary customer response	2019-21	Afternoon of event: email, SMS, social media	No direct reward. Community benefit based on saving retailer cost and lowering prices for all Enova customers.		Reduce unnecessary usage, if it's safe to do so, during specified hours.		Decision to send requests has to balance potential impact, cost of communication and the risk of overloading customer messaging.
Jemena	Power Changers	2017 - 2018		$20/event for participation plus rewards for surveys, as personal gift card or community reward.	Target set as % of average (using CAISO 10 in 10). And DR measured by comparison with control group of similar households.	Included "Learn and earn" education about electricity market	Smart Meter and mobile app	Biased towards users with low consumption. Community incentives less effective than personal but attracted different cohort.
Origin Energy	Origin Spike	2020	Email / SMS	PayPal cash or gift cards for reaching goal set.	CAISO 10 by 10	Gamification: challenge to reduce energy use with small behaviour changes"	Spike platform	
Powershop	Curb Your Power	2017-2021		Evolved to: $10 to curb 10% throughout event AND 1kWh AND min 0.05kWh/h throughout	AEMO baseline: net load for solar households doesn't correlate with temp.	Tested prize draw, "Join the club", Community (charity) rewards.	Smart meter and monitoring app	Message by default has greater impact than moral appeals, education or economic incentives.
United Energy	Summer Saver	2015 - 2020	In app notification and 15-min data updates	"Moved from participation payment to $5/kWh + 50% bonus for all 3 hours		Behaviour suggestions including pre-cooling, close blinds, unplug appliances.	Mobile app & gamification	Near real-time data drove higher engagement.
Zen Ecosystems	RACV and Planet Innovation Behavioural programs	2017-2020	SMS		AEMO Adjusted Baseline. Then modified to "linear baseline"	"Help the Grid" was strong motivator in 17/18 (awareness of blackouts) but far less impactful later.		Costs of the incentives provided not covered by the value of DR delivered. Lack of data access.

4 Experimental Design

The Load-shifting Challenge was designed to inform three key questions regarding demand-response programs that aim to achieve within-day load shifting. These empirical questions (Q) were:

- Q1 Are households more responsive to incentives that encourage load shifting by subsidising daytime electricity use (solar sponges) or by rewarding reductions in evening electricity use (peak shavers)?
- Q2 Are households more responsive to every day (routine) incentives or targeted (ad hoc) incentives?
- Q3 How does household responsiveness vary with the size and type* of incentives? (* monetary and non-monetary)

Moreover, the Load-shifting Challenge examined the economic consequences of the answers to Q1-3. For example, can a financial surplus be generated from these responses, assuming particular daytime and evening price realisations?

We used an RCT experimental design to allow for direct tests of the above questions. The experimental treatment groups are described in this section, followed by the design of the recruitment process and the within-app experience. The next section reports the recruitment outcomes and experimental payouts, with the subsequent section reporting the results from our trial.

4.1 Treatment groups

There were eight treatment groups in the trial. The following features [F] define these treatment groups.

F1 Incentive window

"Into" vs "Away": Groups that have solar sponge incentives that subsidise daytime use between the hours of 12 noon and 3 pm were classed as "Into". Groups that have peak shave incentives that reward reductions in evening use between the hours of 5 pm and 8 pm were classed as "Away" because they encourage moving energy into and away from these time periods, respectively.

F2 Incentive frequency

"Routine" vs "Ad hoc": Groups that received daily incentivised events were classified as "Routine". Groups that had less-than-daily events were classified as "Ad hoc."

F3 Incentive magnitude

The Load-shifting Challenge tested four level of incentives: 5c/kWh, 10c/kWh, 50c/kWh and a non-monetary incentive. For the "Into" groups, participants received a payment for every kWh used during an event window. For the "Away" groups, participants received a payment for every kWh they use below a personalised baseline in an event window.[4] The non-monetary groups received either 0, 1, 2 or 3 stars. These stars had no monetary value, with no stars awarded if the personalised baseline was exceeded, 1 star for consumption below the

[4] Raw baselines were each household's 90th percentile of daily usage between the hours of 5 pm-8 pm from 1 December 2021 through to 14 March 2022. Final baselines were assigned by rounding the raw baselines up to the nearest decile of household raw baselines. The exception was households in the 90th to 99th percentile, which were assigned the 95th percentile value for program budget reasons. We note that most households have peak demand in these summer months, so the baselines were expected to almost always exceed actual consumption during the trial, providing participants in peak shave incentive groups nearly always with a positive opportunity cost from using more energy equal to the incentive magnitude.

baseline, 2 stars for consumption below 50% of the baseline, and 3 stars for consumption below 75% of the baseline.

The eight treatment groups [G], defined by their values of the three features (F1 + F2 + F3), were:

- G0 Control group (received no treatment - the control group was informed that they were not selected for the Load-shifting Challenge trial)
- G1 Away + Routine + 5c/kWh
- G2 Away + Ad hoc + 5c/kWh
- G3 Away + Ad hoc + 10c/kWh
- G4 Away + Ad hoc + 50c/kWh
- G5 Into + Routine + 5c/kWh
- G6 Into + Ad hoc + 5c/kWh
- G7 Into + Ad hoc + 10c/kWh
- G8 Away + Routine + Non-monetary

Comparing the consumption patterns of the different treatment groups with the control group G0 was necessary to identify the average treatment effect for each program, while the comparison across treatment groups spoke to the three empirical questions outlined at the start of this section. For example, comparing G1 to G5, G2 to G6, and G3 to G7 informed Q1 on the load-shifting and subsequent procurement cost impacts of solar sponge program designs relative to comparable peak shave program designs. Comparing G1 to G2, and G5 to G6 informed Q2 on the load-shifting and subsequent procurement cost impacts to daily routine events versus targeted ad hoc events. Finally, comparing G1 to G8, G6 to G7, and G2 to G3 and G4 informed Q3 on how the magnitude of the incentive impacts load shifting and subsequent procurement cost impacts.

4.2 Recruitment and treatment randomisation

The RCT enabled a causal interpretation of the impacts the different program design features had on the consumption profiles of the participants in the Load-shifting Challenge. The recruitment process began with an invitation sent via an app notification and an email to each of Powerpal's 57,979 Victorian customers that had at least 3 months of tenure at the time of the invitation. These users were invited to participate in a challenge that could see them earn up to $50 without giving details regarding the program design and what actions will result in earning rewards.[5] These invitations are documented in Appendix A.

Once registrations closed, we randomly allocated each participant to one of the eight treatment groups or the control group. Those allocated to the control group were not given any further details on the trial and were thanked for their interest by receiving $5. Participants allocated to a treatment group received welcome emails announcing the start date of the program, the functionality of the event page in the app and tips for how to earn rewards. The only content differences across the groups reflected the incentive design features—the incentive window, frequency and magnitude—that define each treatment group. These welcome emails, and the home screens for the trial in the app, are documented in Appendix A. Participants were informed that they would receive their reward payments at the end of the program following the completion of a short exit survey.

[5] All rewards were given to participants in the form of a digital shopping voucher accepted widely among Australian retailers.

4.3 Within-app experience

All participants in the Load-shifting Challenge trial received push notifications on their phone five minutes before each event, alerting them to the start of the event and inviting them to track their progress in real-time in the Powerpal app, which was followed by another notification at the conclusion of the event containing a summary of their event earnings.[6,7] The Powerpal app was an integral part of the Load-shifting Challenge, enabling participants to remind themselves of the parameters and conditions of the program, obtain load-shifting tips, track their progress in real-time as each event progressed, and access previous events.

The within-app experience was custom-designed for each treatment group. The following screenshots pertain to the 5c/kWh routine-peak shave group. The screenshots that applied to the solar sponge and non-monetary groups are provided in Appendix B.

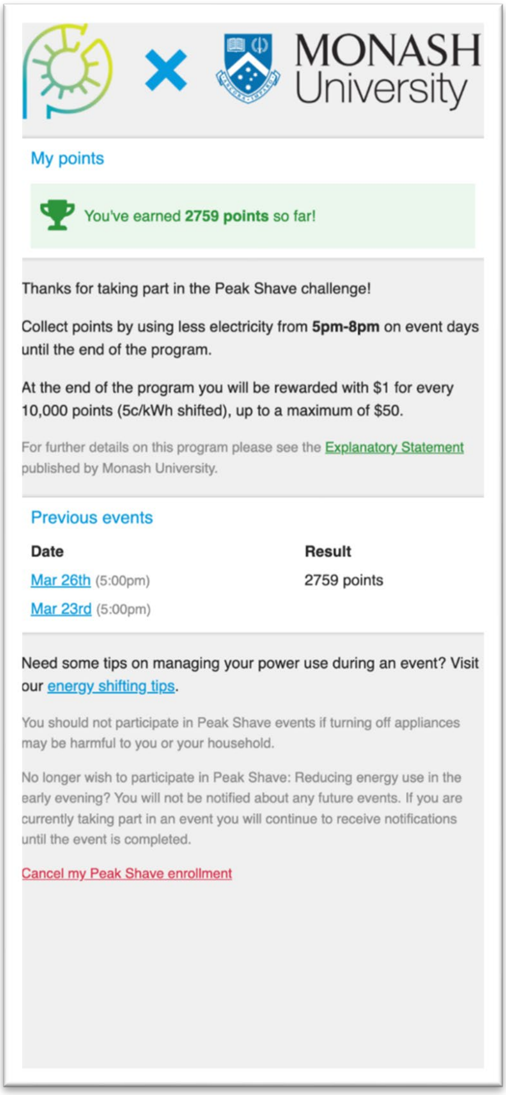


Figure 4-1: Load-shifting Challenge home screen in Powerpal app

---

[6] Participants in the ad hoc groups received an additional notification 24 hours in advance of a scheduled event.

[7] Participants in the monetary treatments differed in the number of reward points they earnt per kWh shifted. All reward points were converted into a dollar amount at the same rate of 10000 to 1, with the conversion rate and equivalent financial reward per kWh shifted clearly shown on the Load-shifting Challenge home screen. Participants in the non-monetary treatment were notified of the number of stars earned during the event.

The Load-shifting Challenge home screen (Figure 4-1) was always accessible in the Powerpal app and provided information on the cumulative reward points earned to date and their monetary value. It also served to remind participants about the key experimental parameters, summarise their performance in previous events and provided warnings as well as links to a load-shifting tips page (provided in Appendix C) and an option to cancel their enrolment at any time.

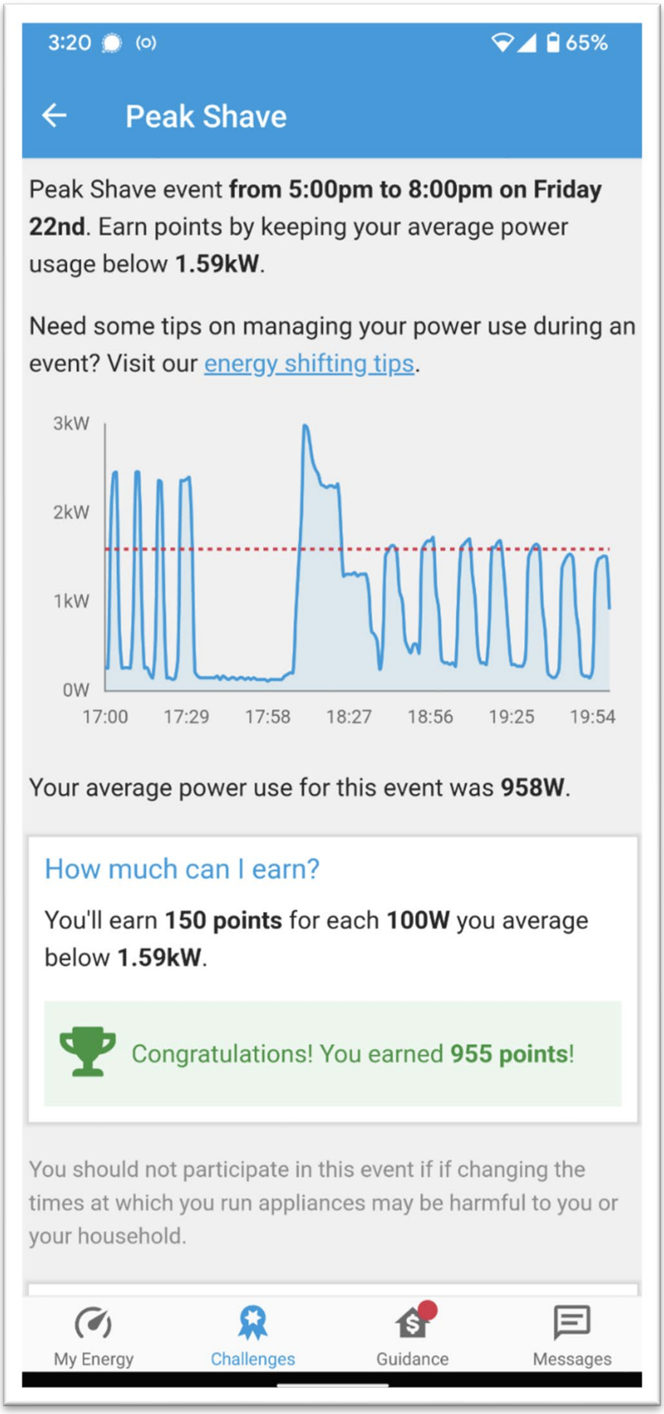


Figure 4-2: Load-shifting Challenge within event screen in Powerpal app

The within-event screen (Figure 4-2) tracked the participants' consumption against their personal baseline in real-time. It also reminded participants of how many points they could earn per kWh shifted and linked to the load-shifting tips page.

5 Data, Recruitment and Events

5.1 Recruitment outcomes

Over 6,000 Powerpal users accepted the invitation to participate in the Load-shifting Challenge, with an 8.9% acceptance rate for users without solar and a 12% acceptance rate for users with solar installations at their home. These users were then randomly assigned into the treatment and control groups, listed in Table 5-1.

Table 5-1: Breakdown of the number of participants

Treatment group	Non-solar users	Solar users
G0: Control	607	809
G1: Away + Routine + 5c/kWh	310	486
G2: Away + Ad hoc + 5c/kWh	310	487
G3: Away + Ad hoc + 10c/kWh	-	489
G4: Away + Ad hoc + 50c/kWh	307	490
G5: Into + Routine + 5c/kWh	308	-
G6: Into + Ad hoc + 5c/kWh	309	-
G7: Into + Ad hoc + 10c/kWh	303	-
G8: Away + Routine + Non-monetary	306	484
Total acceptances	2,760	3,245
Total invited	30,942	27,037

A key feature of the data was that net usage was observed with a floor of 0 kWh. Given that solar panel owners can inject energy into the grid when their home usage is less than their solar panel generation, this resulted in non-trivial data censoring for these users.[8] If the treatment groups were successful at increasing electricity use during daylight hours for solar owners, it was highly likely that we would be unable to empirically observe this fact. For this reason, solar users were concentrated into the "away" treatments. This was because the evening window during the Load-shifting Challenge window coincided with little to no sunlight, and thus net usage was greater than zero, allowing the incentive impacts in these windows to be observed. In contrast, non-solar users had complete data—their net usage was their total usage. As a result, a more complete load-shifting analysis could be performed on this subsample. Non-solar households were excluded from participation in G3 due to ex-ante sample size and statistical power considerations. The inclusion of non-solar households in G3 was not

[8] A data-driven method was used to designate households as solar or non-solar, and this decision was validated by examining self-reported solar installations reported at the time of Powerpal installation. Solar installations were considered to be highly probable if at least twice in the pre-trial period (3 months) a household recorded 0 Wh for 4 consecutive hours; this suggested injections into the grid. 93% of households we classified as having no solar self-reported having no solar. 83% of households we classified as having solar self-reported having solar, however, the self-reporting of solar is an underestimate because it reflects solar installations at the time of installing Powerpal, not at the time of the trial. For this reason, the empirical classification was preferred.

strictly necessary for understanding the sensitivity of users to prices in the "away" treatment, hence their omission was of limited consequence.

To confirm the successful randomisation of our control and treatment groups we observed pre-trial usage patterns across individuals. Successful randomisation would imply no difference in the mean pre-trial usage patterns across groups. Formally, our null and alternative hypotheses were:

- Null: no difference in the pre-trial usage patterns across all hours of the day between the control group and all other groups (or usage patterns are approximately equal)
- Alternative: statistically significant difference in the pre-trial usage patterns between the control group and other groups.

We failed to reject the null hypothesis that pre-trial usage patterns across all hours of the day between the control group and all other groups were equal. This test confirms that randomisation was successful and is outlined in Section 6 after the empirical strategy is introduced. Figure 5-1 displays the hour-level means of our usage data in the pre-trial period (1 December–28 March) across each treatment and control category. Throughout, all usage data is reported in watt-hours, and given the hourly aggregation of the data, this approximates average watts at a point in time within each hour. These figures further highlight the pre-trial similarities between profiles across treatment groups within the solar and non-solar users. The figure also makes clear the differences across solar and non-solar users, with solar users seeing substantially lower net usage on average in the daylight hours.

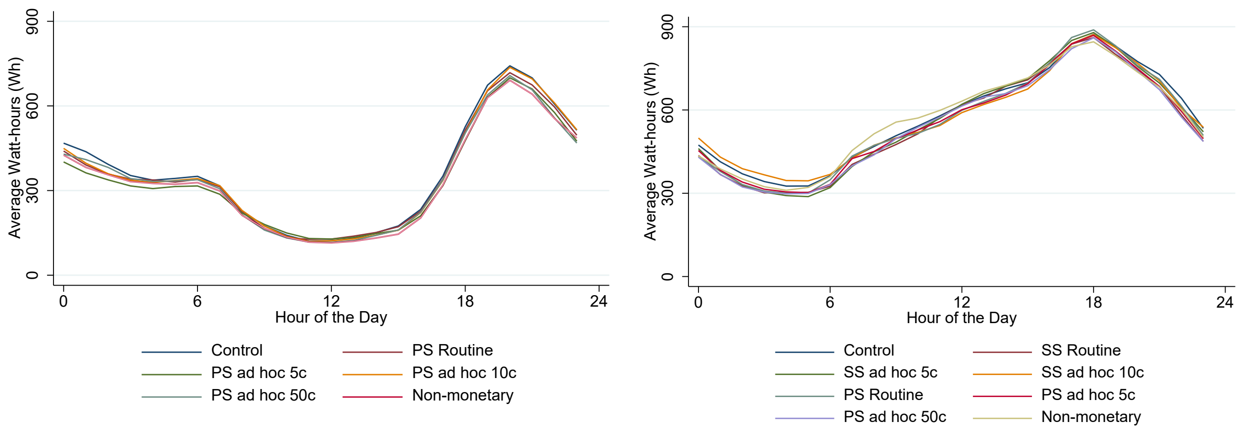


Figure 5-1: Hour-of-day average pre-trial consumption for solar (left) and non-solar (right) households by treatment group

5.2 Events

Events were run every day for the "Routine" groups, G1, G5 and G8. Events for the 5c/kWh and 10c/kWh ad hoc groups occurred 2-3 times per week, with preference given to days that were projected to be sunnier in the daytime given these conditions are usually more conducive to high-value load shifting opportunities. The 50c/kWh group was targeted much less frequently (every 1 to 2 weeks) as a utility or retailer considering implementing Load-shifting Challenge-type programs will trade off larger incentive payments with a lower frequency to target days where peak prices are expected to be abnormally high. Total events and their timing during our trial commencing on 29 March 2022 and ending on 30 June 2022 are summarised in Table 5-2 and Figure 5-2.

Table 5-2: Number of events per treatment group

Group characteristic	Treatment groups	Number of events
Routine	G1, G5, G8	94
Ad hoc + 5c/kWh, 10c/kWh	G2, G3, G6, G7	36
Ad hoc + 50c/kWh	G4	7

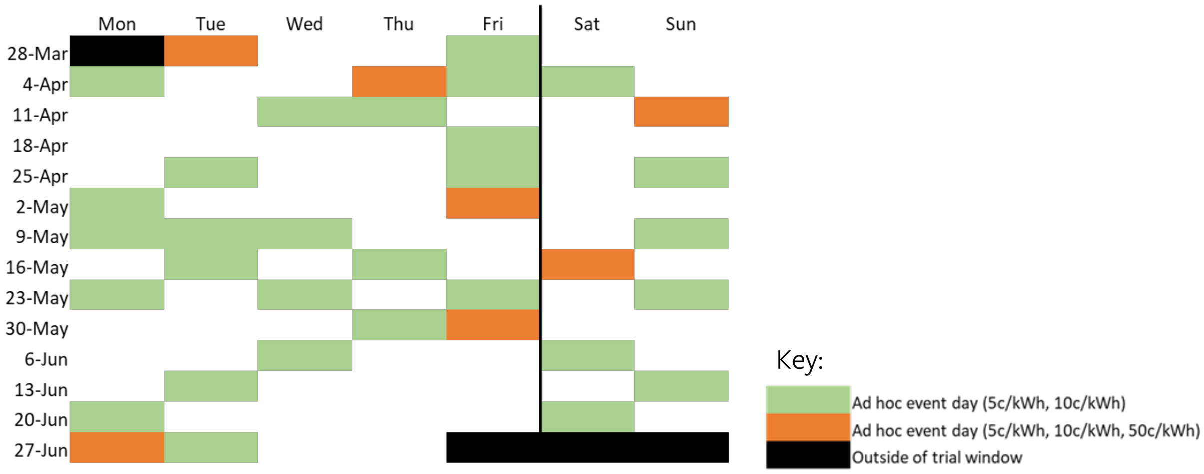


Figure 5-2: Calendar depiction of trial event days

Figure 5-3 shows the half-hourly average wholesale prices on event days for each of the three event types. The ad hoc event days largely had similar price patterns to all days. However, the 50c/kWh event days tended to have lower daytime prices and higher average prices at 5 pm (reflecting that one of the events occurred when wholesale prices spiked). This shows that on average during our trial, the greatest value that could be generated from load shifting and peak shaving occurred on ad hoc 50c event days.

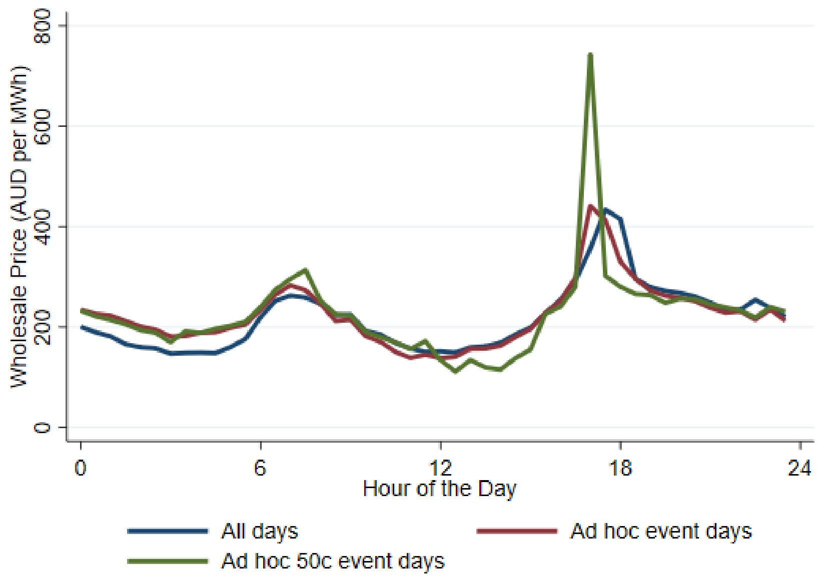


Figure 5-3: Average wholesale electricity prices during trial events

5.3 Experimental payments

By the end of the Load-shifting Challenge trial on 30 June 2022, participants had earned the following rewards summed by treatment group.

Table 5-3: Breakdown of the financial rewards by treatment group

Treatment group	Financial rewards
G0: Control	$9,335
G1: Away + Routine + 5c/kWh	$7,505
G2: Away + Ad hoc + 5c/kWh	$4,745
G3: Away + Ad hoc + 10c/kWh	$3,425
G4: Away + Ad hoc + 50c/kWh	$6,805
G5: Into + Routine + 5c/kWh	$2,460
G6: Into + Ad hoc + 5c/kWh	$1,655
G7: Into + Ad hoc + 10c/kWh	$2,190
G8: Away + Routine + Non-monetary	$3,795

Participants were paid in Prezzee vouchers, that is a voucher with a cash face value that can be redeemed at 100+ major and specialist online retailers, upon completing the exit survey. The implementation of the Prezzee voucher system is low-cost, with transaction costs adding less than 1.5% to the reward amount.[9]

[9] Payments to individual participants have a $5 floor and $50 ceiling.

6 Empirical Strategy

The average impact the treatment assignment had on electricity use for each hour of the day was estimated within the same model. This allowed for post-estimation tests relating to load shifting hypotheses. Two models were specified with the second being a variation on the first, allowing us to estimate the effect of ad hoc event days. The models take a general form similar to that of Andersen et al. (2021), adapted to explicitly take a difference-in-difference format. The models used the following definitions and identifiers, with each observation in the data uniquely identified by the household i and the hour-of-sample t:

$G_{i,g}$ Indicator set to 1 if household i was allocated to treatment group g

S_i Indicator set to 1 if household i had solar panels

$H_{t,h}$ Indicator set to 1 if hour-of-day for hour-of-sample t is equal to h, where h takes values 0 to 23

W_t: Indicator set to 1 if hour-of-sample t was during the trial window (date is 29 March 2022 or later)

A_t: Indicator set to 1 if hour-of-sample t was during an event day for the 5c/kWh and 10c/kWh ad hoc groups (core model specifications do not include the 50c/kWh group)

A'_t: Indicator set to 1 if hour-of-sample t was not during an event day for the 5c/kWh and 10c/kWh ad hoc groups

μ_h, ν_i, η_d: Hour-of-day, household, and day-of-sample fixed effects

$y_{i,t}$: Outcome variable for household i in hour-of-sample t. These are either watt-hours, wholesale energy procurement costs (watt-hours multiplied by the wholesale electricity price in $/Wh), or total energy procurement cost (wholesale costs plus the costs of the incentive payments used in the project). All variables are in levels (without a log transformation) due to the prevalence of zero usage observations (particularly among solar households) and negative wholesale prices.

The first model specification (E-1) took a modified difference-in-difference form, where separate estimates for the treatment effects from assignment to treatment groups were estimated for every hour of the day:

$$y_{i,t} = \mu_h + \nu_i + \eta_d + \sum_{h=0}^{23} \alpha_h H_{t,h} W_t + \sum_{h=0}^{23} \sum_{g} \beta_{g,h} H_{t,h} W_t G_{i,g} + \epsilon_{i,t} \tag{E-1}$$

Here, $\beta_{g,h}$ were the coefficients of interest, with $\epsilon_{i,t}$ independently distributed mean zero error terms. This model was primarily used for partitions of the sample that only include the routine treatment groups—under this partition $\beta_{g,h}$ represents the average treatment effect of assigning participants to treatment group g on outcome variable y during hour-of-day h. Recall that the control group was formed of people willing to participate in the Load-shifting Challenge but who did not receive any further details on the trial after being notified it was oversubscribed, making these average treatment effects within the set of households that were open to participating in the trial. $(\beta_{g,h} - \beta_{g',h})$ corresponds to the average treatment effect of assigning participants to treatment group g relative to a counterfactual assignment to treatment group g' on outcome variable y during hour-of-day h. If ad hoc groups are included in this model, then $\beta_{g,h}$ represents the average treatment effect for assignment to group g over both non-event and event days. The second model specification (E-2) separates out the average treatment effects during non-event and event days for ad hoc groups. We used Ordinary Least Squares to estimate each model, with standard errors clustered at the household level reported and used to conduct inference.

The following 6 Wald tests were used to inform the impacts of assignment to the various treatment groups g:

Aggregate impact: $\sum_{h=0}^{23} \beta_{g,h} = 0$ (T-1)

Daytime target hours: $\sum_{h=12}^{14} \beta_{g,h} = 0$ (T-2)

Evening target hours: $\sum_{h=17}^{19} \beta_{g,h} = 0$ (T-3)

Load shifting – daytime immediate: $\beta_{g,12} - \beta_{g,11} = 0; \beta_{g,15} - \beta_{g,14} = 0$ (T-4)

Load shifting – evening immediate: $\beta_{g,17} - \beta_{g,16} = 0; \beta_{g,20} - \beta_{g,19} = 0$ (T-5)

Load shifting – incentive windows: $\sum_{h=12}^{14} \beta_{g,h} - \sum_{h=17}^{19} \beta_{g,h} = 0$ (T-6)

Note that tests were able to be modified to compare two treatment groups, for example (T-1), the aggregate impact test, becomes $\sum_{h=0}^{23} \beta_{g,h} = \sum_{h=0}^{23} \beta_{g',h}$.

The second model specification (E-2) took the form

$$y_{i,t} = \mu_h + \nu_i + \eta_d + \sum_{h=1}^{23} \alpha_h H_{t,h} A_{g,t} W_t + \sum_{h=1}^{23} \alpha_h^{'} H_{t,h} A_{g,t}^{'} W_t + \sum_{h=0}^{23} \sum_{g} \gamma_{g,h} H_{t,h} A_{g,t} G_{i,g} + \sum_{h=0}^{23} \sum_{g} \gamma_{g,h}^{'} H_{t,h} A_{g,t}^{'} G_{i,g} + \epsilon_{i,t} \tag{E-2}$$

Here, $\gamma_{g,h}$and $\gamma'_{g,h}$were the coefficients of interest. This specification allowed for additional testing of hypotheses related to habit formation. The aforementioned tests relating to the first specification can also be conducted with interpretations being refined to the average impact of the treatment group assignment on outcomes on ad hoc event days ($\gamma_{g,h}$) or ad hoc non-event days ($\gamma'_{g,h}$). For brevity, we list tests that related to habit formation with peak shaving during the incentive windows.

Habit formation – routine and ad hoc comparison: $\sum_{h=17}^{19} \gamma_{g,h} = \sum_{h=17}^{19} \gamma_{g',h}$ (T-7)

Habit formation – control and ad hoc comparison: $\sum_{h=17}^{19} \gamma'_{g,h} = 0$ (T-8)

Habit formation – ad hoc event / non-event comparison: $\sum_{h=17}^{19} \gamma'_{g,h} = \sum_{h=17}^{19} \gamma_{g,h}$ (T-9)

Any habit formation resulting from the Load-shifting Challenge was analysed by examining the electricity usage of the routine and/or ad hoc treatment groups. For example, households receiving routine incentives to load shift may experience lower barriers to shift energy on event days relative to those receiving ad hoc incentives because they are more practiced and have long-run certainty that events are called each day. Alternatively, if there is a stock of attention and energy that households can devote to load shifting, then the routine group will deplete this stock more than the ad hoc group, and responses on event days will be greater for the ad hoc participants. Failure to reject the routine and ad hoc comparison test (T-7) is consistent with our treatments generating no habit formation and not depleting a stock of attention and energy that households can devote to load shifting.

However, if habit formation occurred among participants receiving ad hoc incentives, then we might expect to see that they load shifted on days when there is no financial incentive to do so. We should fail to reject the null of the control and ad hoc comparison test (T-8) if our treatments did not elicit any habit formation among our participants, all else being equal. However, we acknowledge that the inference we can draw from this test is perhaps weaker than the routine and ad hoc comparison tests: A rejection of the null here could occur for reasons unrelated to habit formation. For example, observing differences between the control group and an ad hoc group on non-event days could be due to the initial and ongoing communication received by participants in the ad hoc group that motivated load shifting in addition to the financial incentives offered on event days (whereas this messaging is common to both ad hoc and routine groups being tested in the comparison tests). This motivated the final test specification. The ad hoc event/non-event comparison test (T-9) examined whether treatment effects were equivalent on event and non-event days, with a failure to reject the null suggesting responses were not driven by the financial incentives and/or there are habit formation impacts from the Load-shifting Challenge that resulted in equivalent load shifting on non-event and event days.

6.1 Randomisation

Identification of treatment impacts was predicated on the randomised assignment of users to treatment groups. We briefly describe the statistical support confirming a successful randomisation among the users that completed the trial and enter the analysis sample.

Firstly, we conducted simple comparison of means tests for the average usage for each participant across different treatment groups. The mean hourly usage for the control group participants pre-trial was 572.8 Wh for non-solar users and 357.5 Wh for solar users. We failed to reject the equivalence of means at a 5% level for any treatment group compared to the relevant solar/non-solar control group, or when compared to all relevant solar/non-solar participants outside of their treatment group.

Secondly, we tested for equivalent load profile shapes by modifying equation (E-1) by collapsing the data to contain the average watt-hours for each hour of the day for each participant ($y_{i,h}$), and define indicator variable H_h, which is set to 1 if hour-of-day is h, that is, there are 24 observations per participant. The model below is estimated with standard error clustered at the participant-level.

$$y_{i,h} = \sum_{h=1}^{23}[\alpha_h^S H_h S_i + \alpha_h^{NS} H_h(1 - S_i)] + \sum_{h=0}^{23}\sum_{g}[\beta_{g,h}^S H_h S_i G_{i,g} + \beta_{g,h}^{NS} H_h(1 - S_i)G_{i,g}] + \epsilon_{i,h}$$

For each treatment group we tested $\beta_{g,0}^S = \beta_{g,1}^S = \cdots = \beta_{g,23}^S = 0$ and $\beta_{g,0}^{NS} = \beta_{g,1}^{NS} = \cdots = \beta_{g,23}^{NS} = 0$, where the superscripts *S* and *NS* refer, respectively, to solar and non-solar users. We failed to reject each test for each treatment group at a 5% level and concluded that we do not have evidence that suggests any treatment group has a statistically significant deviation in their average hour-to-hour load profile pre-trial. We consequently discuss the drivers of differences we identify in load profiles as impacts of the treatments and do not attribute these differences to any inherent pre-existing differences in the electricity use habits of participants in each treatment group.

7 Results

The results are structured into two main sections, electricity use and electricity costs, followed by some insights from the end-of-trial survey. The electricity use section considers the impacts of the different incentive types on load shifting — "into" (solar sponge) versus "away" (peak shave), "routine" versus "ad hoc"— and the magnitude of the incentives offered (5c/kWh, 10c/kWh and 50c/kWh). We begin by highlighting the different responses by solar and non-solar households to the same routine incentive designs. We then contrast responses to these routine incentives with responses to ad hoc incentives. The electricity costs section then examines how wholesale procurement and program costs differ across the treatments. Finally, the survey section supplements the usage results by describing the self-reported characteristics and stated experiences of participants in the trial.

7.1 Treatment effects on household electricity use

7.1.1 Solar and non-solar households: same incentives, different responses

The treatment groups G1 (routine, peak shave, 5c/kWh incentive) and G8 (routine, peak shave, non-monetary incentive) provided a clear starting point for examining how responses to treatments can differ across participant characteristics.

Given the large number of coefficients estimated in this model (24 hours multiplied by G treatment groups in addition to hour-of-day and day-of-sample fixed effects), we summarised subsets of the coefficients and standard errors in a figure. The black line provides the average hourly usage for the control group during the trial window purely for descriptive purposes and to act as a baseline for reporting the departures of each treatment group from the control group.[10] In reporting estimates from equation (E-1), the red line adds $\hat{\beta}_{g,h}$ to the control group series; this reflects the estimated average usage for households in group g in hour h. A pointwise 95% confidence interval brackets this series, adding $\hat{\beta}_{g,h} \pm 1.96.se(\hat{\beta}_{g,h})$ to the control group series. For any given hour of the day where the control group series lies within the 95% confidence interval, we failed to reject the null hypothesis that the mean consumption for the control group and group g are equal for a test with size 5%, and conversely, we rejected this null hypothesis when the control group series lies outside the 95% confidence interval.

Figure 7-1 displays a sub-set of the coefficient estimates of equation (E-1) for solar and non-solar households relating to allocation into the routine, peak shave, 5c/kWh incentive group. The first subfigure reports estimates for the sample of solar households only, and the second subfigure reports estimates for the sample of non-solar households only.[11] We detected statistically significant impacts among the solar households in the 5 pm-8 pm peak shave window; on average, participants used 192 Wh (9%) less energy at that time than the solar control group. However, in the second sub-figure we did not identify any impact for the equivalent treatment among non-solar users for any hour of day.

[10] The control group series for model (E-1) is the average of the estimated day fixed-effects, $\hat{\eta}_d$ during the trial window added to the relevant hour-fixed effects, $\hat{\mu}_h$ and $\hat{\alpha}_h$.

[11] In practice there are two groups g per experimental group in equation (E-1), one for solar households and another for non-solar households. Further, the hourly fixed effects are also interacted with solar/non-solar status.

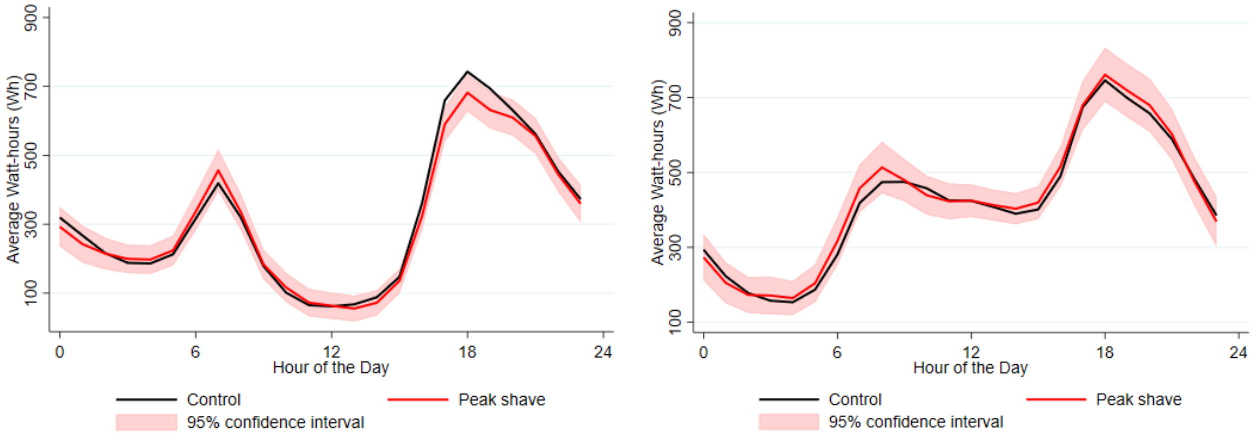


Figure 7-1: Average treatment effects for routine, peak shave, 5c/kWh incentives; Solar users (left) and Non-solar users (right)

Figure 7-2 similarly displays a sub-set of the coefficient estimates of equation (E-1) but for the routine, peak shave, non-monetary incentive groups. We observed similar results to the paid incentive group—statistically significant impacts among the solar households in the 5 pm-8 pm peak shave window; average reduced energy use by participants of 163 Wh (8%) at that time than the control group—but we did not identify any impact for the equivalent treatment among non-solar users for any hour of day. Finally, we failed to reject tests for equivalent treatment effects for the 5 pm-8 pm window for solar non-monetary and 5c/kWh incentives, suggesting that similar reductions in energy use can be achieved without a monetary incentive for this class of household.[12]

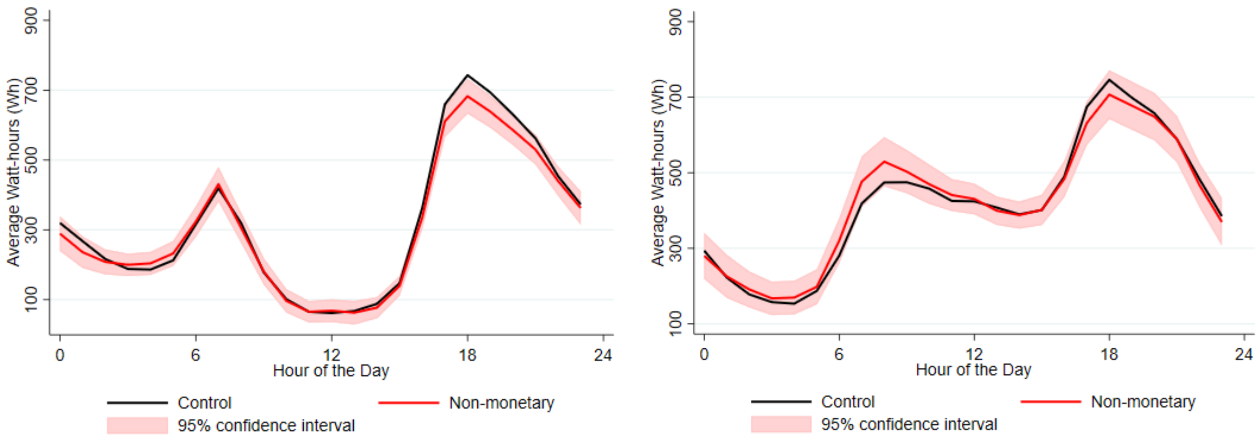


Figure 7-2: Average treatment effects for routine, peak shave, non-monetary incentives; Solar users (left) and Non-solar users (right)

The non-solar participants displayed no evidence to suggest that they were responsive to monetary or non-monetary incentives to use less energy in the 5 pm-8 pm peak shave window. However, there is evidence that they were responsive to incentives to use more energy in the 12 pm-3 pm solar sponge window, demonstrating that this class of customer is sensitive to the structure of incentive offered. Figure 7-3 reports estimates of equation (E-1) relating to the solar sponge incentives for the subset of non-solar customers allocated to the

[12] The p-values for Wald tests for the sum of the relevant 5 pm, 6 pm and 7 pm coefficients are: solar, monetary incentives (relative to the solar control group), 0.013; non-solar monetary incentives (relative to the non-solar control group), 0.69; solar, non-monetary incentives (relative to the solar control group), 0.017; non-solar, non-monetary incentives (relative to the non-solar control group), 0.25. Finally, the test for the equivalent treatment effect for the monetary and non-monetary groups of solar customers returns a p-value of 0.72.

control or routine treatment groups. There are two striking features: First, there was a noticeable change in trajectory of the load shape at 12 pm, when the incentive window starts, suggesting that the incentive is motivating more usage at that time. Second, there was less energy use for most of the peak energy hours relative to the control. This suggests that a more effective incentive design to encourage the shaving of energy use by non-solar households in the evening is not to reward *less* energy use at that time but to instead encourage *more* energy use at an earlier time. On average, households facing this incentive design used 85 Wh (6%) more energy during the 12 pm-3 pm incentivised solar sponge window, and 188 Wh (8%) less energy in the 5 pm-8 pm peak window that was not subject to incentive payments.

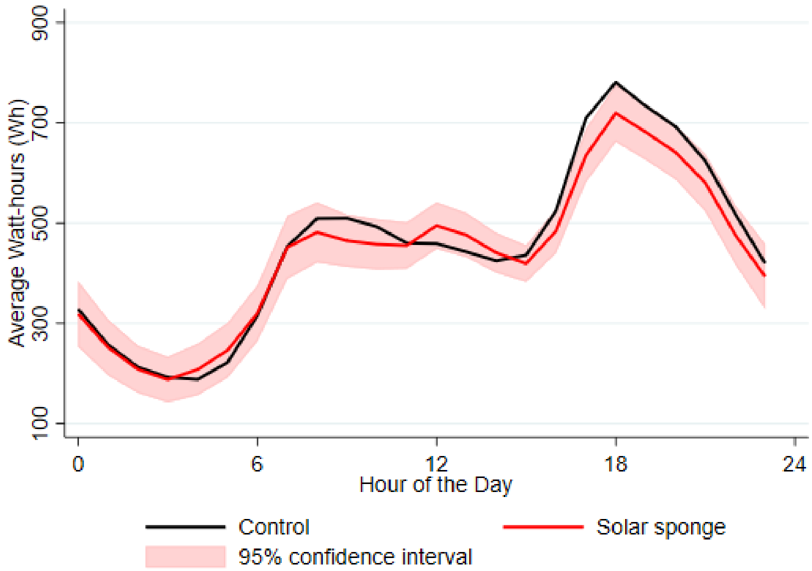


Figure 7-3: Average treatment effects for routine, solar sponge, 5c/kWh incentives: Non-solar users

Although there appeared to be an overall conservation of energy for non-solar households allocated to the solar sponge incentive group, the first entry in Table 7-1 shows that we fail to reject the null hypothesis, that overall energy consumption differs from the control group, at a 5% test size (T-1). However, we find evidence of load shifting: The third entry (T-3) of Column 1 shows that we identify peak shaving in the evening window relative to the control group, with the fourth entry (T-6) providing strong evidence that solar sponge incentives significantly reduced the disparity between middle-of-the-day and evening consumption. Similarly, the fifth entry (T-4) shows that we can reject that the consumption trajectory entering and exiting the daytime incentive windows was the same as for the control group. That is, we have evidence to suggest that the incentives resulted in some households delaying their 11 am usage until 12 pm or bringing forward 3 pm usage to 2 pm. Similar evidence exists for consumption shifts into the hours immediately prior and after the evening target windows.

The second and third columns of this table demonstrate that no consumption responses or load-shifting responses were identified for non-solar households subject to monetary or non-monetary incentives to reduce usage in the evening hours.[13]

[13] Note that this testing regime was not repeated for the solar experimental groups due to the binding daytime data censoring for these households making identification of load shifting unviable compared to the peak shaving results already discussed.

Table 7-1: Wald test p-values for tests of treatment effects relating to usage and load shifting: Non-solar households, routine incentives

			Solar sponge	Peak shave	Non-monetary
(T-1)	Aggregate	$\sum_{h=0}^{23} \beta_{g,h} = 0$	0.324	0.661	0.816
(T-2)	Daytime target hours	$\sum_{h=12}^{14} \beta_{g,h} = 0$	0.177	0.757	0.949
(T-3)	Evening target hours	$\sum_{h=17}^{19} \beta_{g,h} = 0$	0.019	0.688	0.256
(T-6)	*Load shifting – incentive windows*	$\sum_{h=12}^{14} \beta_{g,h} - \sum_{h=17}^{19} \beta_{g,h} = 0$	0.000	0.793	0.236
(T-4)	*Load shifting – daytime immediate*	$\beta_{g,12} - \beta_{g,11} = 0;$ $\beta_{g,15} - \beta_{g,14} = 0$	0.000	0.830	0.615
(T-5)	*Load shifting – evening immediate*	$\beta_{g,17} - \beta_{g,16} = 0;$ $\beta_{g,20} - \beta_{g,19} = 0$	0.048	0.444	0.101

Summary: Non-solar households appear more sensitive to incentive design than solar households.

Monetary and non-monetary peak shave incentives did not elicit a response for non-solar households. Solar households responded to either form of incentive, which suggests that this class of customer is motivated to act by being in a program, but without a detectable degree of price elasticity from the 0c/kWh to 5c/kWh incentive range.

Non-solar households did, however, respond to monetary solar sponge incentives where they were paid to use more energy in the middle of the day in a very encouraging manner for policy-makers and retailers that aim to reduce the average cost of energy for these customers: They shifted energy use into the middle-of-day incentive window *and* reduced their energy consumption during the evening peak despite not receiving a direct incentive for this latter action.

7.1.2 Routine and ad hoc incentives: opportunity for demand flexibility

The treatment groups G1 (routine, peak shave, 5c/kWh incentive) and G2 and G3 (ad hoc, peak shave, 5c/kWh and 10c/kWh incentives) were suitable for examining how responses to ad hoc incentives differ from routine incentives for solar households. Given non-solar households did not respond to peak shave incentives, we limited the study for these households to how responses differed between ad hoc and routine solar sponge incentives (G5, G6 and G7).[14] This section reports estimates of equation (E-2).

The average treatment effects on ad hoc event days (the days coloured either green or orange in Figure 5-2) were considered first for solar households that were allocated to the routine peak shave group (5c/kWh) and the ad hoc peak shave groups (5c/kWh and 10c/kWh). The 5c/kWh and 10c/kWh ad hoc groups were combined throughout this section for additional power because we did not identify a statistically significant difference

[14] We did not find any statistically detectable treatment effects for non-solar households allocated to the ad hoc peak shave incentive group (similar to the routine peak shave groups discussed in the previous section) and subsequently do not report any further outcomes for this group in our analysis.

across these two payment levels.[15] It was found that both ad hoc and routine groups reduced their usage relative to the control group in the 5 pm-8 pm peak shave window, depicted in Figure 7-4. Testing failed to reject that these treatment effects were equal (i.e. there was no evidence that the consumption profiles of ad hoc and routine groups differed).[16] This suggests that there may be no routine formation benefits from every day incentives that drive bigger changes to consumption profiles, nor may it be problematic for households to respond to ad hoc events with only a day of notice relative to households that have events every day.

Finally, those with ad hoc incentives were tested to see if they also altered their consumption profile relative to the control group on non-event days. No evidence was found that there were different means over the 5 pm, 6 pm and 7 pm peak shave window on non-event days relative to the control, suggesting this was a flexible response to an ad hoc incentive and that the impacts did not spill over to non-event days via a change in habit or some other form of consumption response.[17]

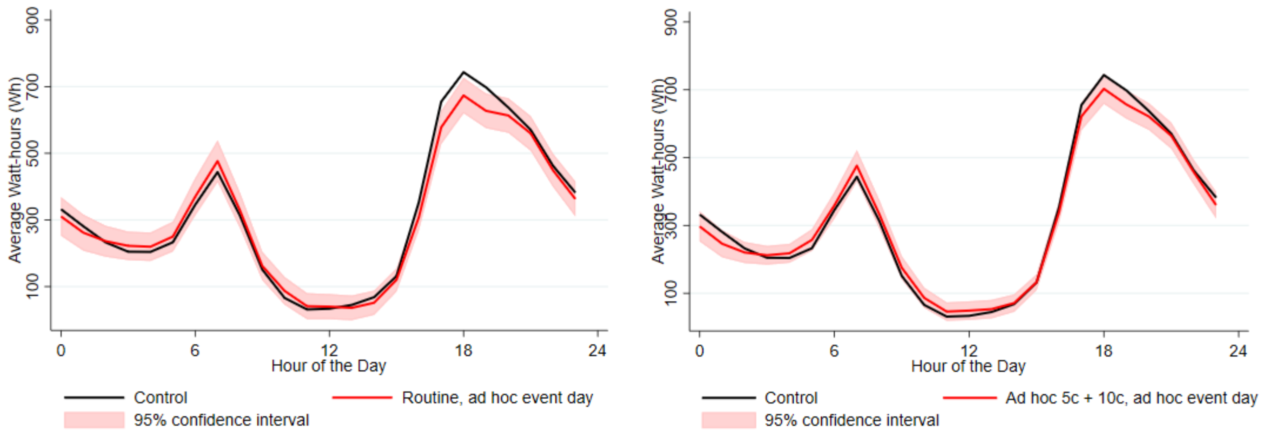


Figure 7-4: Average treatment effects on ad hoc event days, solar households, peak shave incentives; Routine incentives (left) and Ad hoc incentives (right)

Next, non-solar household responses to ad hoc solar sponge incentives on event days (Figure 7-5) were examined. They produced a similar set of findings to the solar ad hoc results, with no evidence of differences between those presented with routine or ad hoc incentives on event days in either the 12 pm-3 pm or 5 pm-8 pm windows.[18] Again, this result supports a claim that there are no benefits from routine formation on driving bigger changes to consumption profiles, nor is it problematic for households to respond to ad hoc events with only a day of notice relative to households that have events every day. Further, it found no evidence that there were different mean levels of consumption for those in the ad hoc incentive groups on

[15] We failed to reject a joint test for equivalent treatment effects in the solar sponge (12 pm-3 pm) and peak shave (5 pm-8 pm) windows for a test of 5% size when we allow the 5c and 10c groups to enter as separate experimental groups. That is, we tested $\sum_{h=12}^{14} \gamma_{g,h} = \sum_{h=12}^{14} \gamma_{g',h}$ and $\sum_{h=17}^{19} \gamma_{g,h} = \sum_{h=17}^{19} \gamma_{g',h}$ where g is the ad hoc 5c peak shave group and g' the ad hoc 10c peak shave group for solar households and fail to reject with a p-value of 0.96. We also tested $\sum_{h=12}^{14} \gamma_{g,h} = \sum_{h=12}^{14} \gamma_{g',h}$ and $\sum_{h=17}^{19} \gamma_{g,h} = \sum_{h=17}^{19} \gamma_{g',h}$ where g is the ad hoc 5c solar sponge group g' the ad hoc 10c solar sponge group for solar households and fail to reject with a p-value of 0.42.

[16] Formally, we reject a 5% size test that $\sum_{h=17}^{19} \gamma_{g,h} = \sum_{h=17}^{19} \gamma_{g',h}$ where g is the ad hoc 5 and 10c peak shave group and g' the routine peak shave group for solar households with a p-value of 0.17.

[17] Formally, we reject a 5% size test that $\sum_{h=17}^{19} \gamma'_{g,h} = 0$ where g is the ad hoc 5 and 10c peak shave group for solar households with a p-value of 0.54.

[18] We failed to reject a joint test for equivalent treatment effects in the solar sponge (12 pm-3 pm) and peak shave (5 pm-8 pm) windows for a test of 5% size. That is, we tested $\sum_{h=12}^{14} \gamma_{g,h} = \sum_{h=12}^{14} \gamma_{g',h}$ and $\sum_{h=17}^{19} \gamma_{g,h} = \sum_{h=17}^{19} \gamma_{g',h}$ where g is the ad hoc 5 and 10c solar sponge groups and g' the routine solar sponge group for non-solar households and fail to reject with a p-value of 0.11.

non-event days when compared to the control group in the key windows of 12 pm, 1 pm, 2 pm and 5 pm, 6 pm, 7 pm, suggesting they can provide a flexible response to an ad hoc incentive with no habit changes or other forms of consumption response spilling over to non-event days.[19]

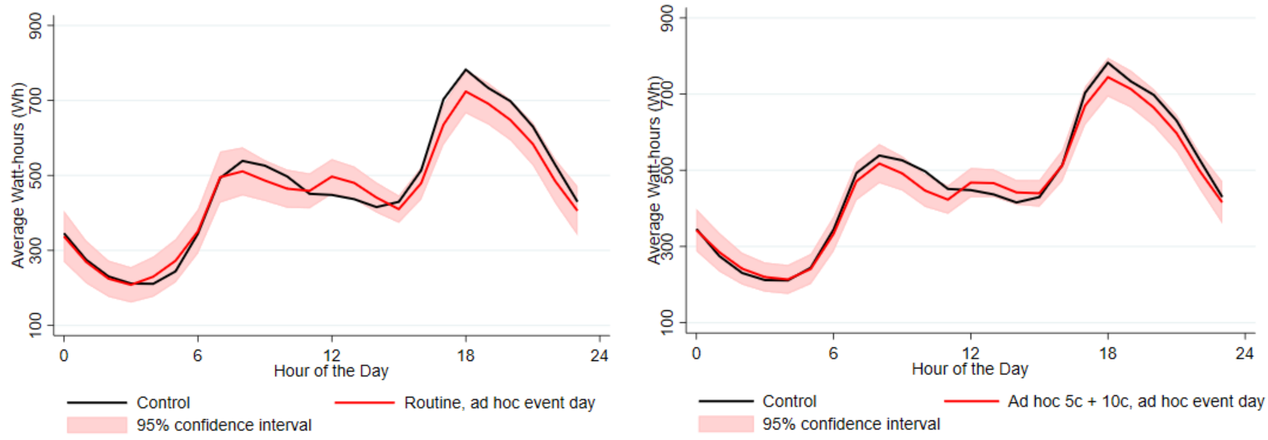


Figure 7-5: Average treatment effects on ad hoc event days, non-solar households, solar sponge incentives; Routine incentives (left) and Ad hoc incentives (right)

Finally, equation (E-2) was re-specified and estimated separately for the smaller set of events that occur for the 50c/kWh peak shave incentive group.[20] Households assigned to the 50c/kWh group received a much larger payment than those in other peak shave groups for reducing their energy use but were eligible for fewer events. However, the experimental design still had all groups presented with incentives on these few event days, allowing for a comparison of responsiveness on these days. The model was estimated separately for solar and non-solar households, and for brevity Figure 7-6 only displays the estimates for those in the 50c/kWh incentive group and not the routine or ad hoc 5c/kWh or 10c/kWh groups on these specific event days. It can be seen that both solar and non-solar households were responsive on these event days, reducing their energy use in the evening relative to the control. This is particularly striking for the non-solar households because no impacts were detected from any other form of peak shave incentive (see prior section). In contrast, the 50c/kWh incentive specification suggests that non-solar households will be responsive to peak shave incentives provided they are large enough. Further, their consumption profile displayed a shaving of energy use in the evening without any other within-day differences, meaning all the responsiveness was from conservation of energy and not load shifting to earlier in the day.

[19] That is, we tested $\sum_{h=12}^{14} \gamma'_{g,h} = 0$ and $\sum_{h=17}^{19} \gamma'_{g,h} = 0$ where g is the ad hoc 5 and 10c solar sponge groups for non-solar households and return a p-value of 0.31.

[20] That is, we respecify A_t and A'_t to reflect ad hoc 50c event days and non-event days.

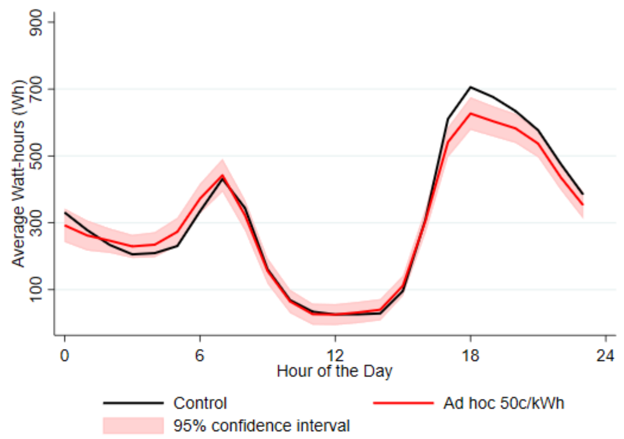


Figure 7-6: Average treatment effects on ad hoc 50c/kWh event days, 50c/kWh peak shave incentives. Solar households (left) and non-solar households (right)

Summary: Ad hoc incentives changed consumption profiles on event days, demonstrating the potential to generate flexible demand response.

The routine incentive treatments we offered suggest that time-of-use-style tariffs and incentives can change load shapes and lower the average cost of supplying households. However, these tariffs do not provide demand response because they do not promote flexibility with respect to real-time conditions. On the other hand, ad hoc events can be called when market or network conditions are such that changing load profiles are forecast to be highly valuable on specific days. These results demonstrate that meaningful consumption responses can be generated on ad hoc event days, and of similar size to those offered by routine incentives.

Finally, it was re-affirmed that non-solar households were more sensitive to incentive design, with average consumption in the evening decreasing on event days when offered 50c/kWh. Although they clearly were more responsive to solar sponge incentives compared to peak shave incentives when payments were low (5c/kWh), they did respond to ad hoc peak shave incentives when the payments were large enough (50c/kWh), further indicating the possibility for them to participate in flexible demand response.

7.2 Electricity costs

The core results from this report relate to electricity use. Given the demonstrated ability to alter load profiles with our incentive designs, price series can be envisaged whereby these shifts create economic value and a surplus to be shared between a household and their retailer. There are two parts to this claim. Firstly, the incentive design ensures that households cannot be worse off by participating: participation is conditional on households opting in and they will accumulate money in each treatment group (or receive a participation payment in the non-monetary incentive group) even without changing their behaviour. Secondly, as long as households move energy away from high-cost times of day and toward low-cost times of day, social costs (the sum of costs from electricity production, carbon emissions and other network costs that arise from using electricity) will decrease, meaning that economic value and/or a retailer surplus can be created without harming the household.

In this section we briefly demonstrate how the changes in consumption profiles can lower the average wholesale energy procurement costs for customers. We estimate equation (E-1) separately for solar and non-solar households, with the dependent variable either the wholesale energy procurements costs (watt-hours multiplied by the wholesale electricity price in \$/Wh), or total energy procurement cost (energy procurement costs plus the costs of the incentive payments used in the Load-shifting Challenge), using the actual wholesale prices observed during our trial. [21] The average daily cost was derived from the estimates of the model. For the control group, this was $\sum_{h=0}^{23}(\tilde{\eta} + \hat{\mu}_h + \hat{\alpha}_h)$, where $\tilde{\eta}$ is the average day-of-sample fixed effect estimated in the model over the 28 March – 30 June trial window. For treatment group g, $\sum_{h=0}^{23} \hat{\beta}_{g,h}$ was added to the control group estimate. Wald tests outlined in (T-1) for $\sum_{h=0}^{23} \hat{\beta}_{g,h} = 0$ were conducted for each treatment group g, with the null hypothesis stating that the average daily wholesale procurement cost per kWh does not differ between the control group and treatment group g.

Table 7-2 shows the average wholesale procurement cost estimates for solar households (acknowledging the imprecision of daytime usage data), and Table 7-3 shows the average wholesale procurement cost estimates for non-solar households. Below each estimate is the p-value of the Wald test for equivalence with the control group (T-1). The final column aggregates the estimates to the 90-day trial window. First examining solar households, we find that all treatment groups had lower wholesale procurement costs than the control group, suggesting that the different consumption profiles saved procurement costs. However, only the non-monetary group and 50c/kWh ad hoc group had savings that were large enough to reject equivalence for a test size 5% or 10% (p-values of 0.05 and 0.06). When including the incentive costs into these estimates, only the non-monetary group remained significantly lower than the control group. Although solar users were responsive to all incentive designs, during our window we only have significant statistical evidence that the non-monetary group was cost-effective.

Table 7-2: Estimates of wholesale procurement costs and program costs for solar households

	Wholesale procurement cost ($, daily average)	Wholesale procurement and program cost ($, daily average)	Wholesale procurement and program cost ($, 90-day program)
Control	$2.91	$2.91	$261.90
Routine peak shave (5c/kWh)	$2.83 *(0.44)*	$2.92 *(0.95)*	$262.80
Ad hoc peak shave (5 + 10c/kWh)	$2.81 *(0.27)*	$2.86 *(0.58)*	$257.40
Ad hoc peak shave (50c/kWh)	$2.72 *(0.06)*	$2.80 *(0.25)*	$252.00
Non-monetary	$2.72 *(0.05)*	$2.72 *(0.05)*	$249.80 ($244.80 + $5)

Note: p-values for Wald test for equivalence to control group listed in parentheses. Program cost for non-monetary group includes the $5 participation payment—their payment was not dependent on their electricity use.

[21] The incentive payments were p c/kWh $* \, kWh$ for participants receiving a solar sponge payment of p c/kWh in a given hour, or max (0,p c/kWh $* \, (B - kWh)$) for participants receiving a peak shave payment of p c/kWh for usage below their baseline B in a given hour.

Table 7-3: Estimates of wholesale procurement costs and program costs for non-solar households

	Wholesale procurement cost ($, daily average)	Wholesale procurement and program cost ($, daily average)	Wholesale procurement and program cost ($, 90-day program)
Control	$3.51	$3.51	$315.90
Routine peak shave (5c/kWh)	$3.49 (0.91)	$3.64 (0.50)	$327.60
Ad hoc peak shave (5c/kWh)	$3.44 (0.71)	$3.50 (0.98)	$315.00
Ad hoc peak shave (50c/kWh)	$3.40 (0.45)	$3.52 (0.94)	$316.80
Routine solar sponge (5c/kWh)	$3.18 (0.01)	$3.27 (0.10)	$294.30
Ad hoc solar sponge (5 + 10c/kWh)	$3.30 (0.12)	$3.35 (0.25)	$301.50
Non-monetary	$3.46 (0.76)	$3.46 (0.76)	$316.40 ($311.40 + $5)

Note: p-values for Wald test for equivalence to control group listed in parentheses. Program cost for non-monetary group includes the $5 participation payment—their payment was not dependent on their electricity use.

Conversely, for the non-solar groups program costs were not significantly different from the control for any of the peak shaving or non-monetary treatments. This was not surprising given that there were no observed consumption profile changes for non-solar households across many of the incentive designs. However, evidence was found that the routine solar sponge treatment drove lower wholesale procurement costs (p-value 0.01), with savings still meaningful when including the program costs. It was estimated that the wholesale procurement costs of these customers were $29 (9%) lower than the control group over 90 days, or $22 (7%) when including the program payments.

Summary: Many of the treatments changed the consumption profiles of participants over the duration of the trial. Therefore, wholesale price conditions and patterns can be envisaged whereby price-reflective tariff changes could generate economic value that could be shared between households and retailers. It was demonstrated that, under prevailing wholesale market conditions during this three-month trial, some treatments decreased wholesale procurement costs, providing a demonstration of the potential for these programs to create value.

7.3 Survey insights

Upon completion of the trial window, participants (excluding those in the control group) were invited to complete a survey that aimed to understand their experience in the trial and what, if anything, motivated or facilitated their load-shifting efforts. Presented are some results that highlight how households' behaviours changed conditional on their incentive group, how motivations differed between solar and non-solar households, potential explanations for how the Load-shifting Challenge impacted solar and non-solar users differently, and how households use of load timers changed during the challenge.

First, the focus was on the non-solar households because they were allocated to all treatment features (i.e. solar sponge, peak shave, and non-monetary). 60%, 67% and 57% of those allocated to solar sponge, peak shave, and non-monetary incentive groups responded to the survey, perhaps reflecting a slightly higher average reward for those allocated to a peak shave group and thus a greater incentive to respond to the survey and claim their reward. However, of these respondents, 95%, 95% and 96% stated that if the Load-shifting Challenge were run again, they would choose to participate, suggesting the participant experience was largely positive.[22]

Figure 7-7 displays responses across solar sponge, peak shave, and non-monetary incentive groups about whether non-solar participants attempted to use more energy or less energy in the day or evening, and whether they tried harder to be home or away from the home during the day or evening. It was seen that those given financial peak shave incentives responded similarly to those receiving non-monetary incentives, with a slightly higher number of those facing a financial incentive (89% vs 82%) stating they made an effort to decrease their evening energy use.[23] However, those offered solar sponge incentives were much more likely to have tried to be home during the day (30% vs 12%) and more likely to state that they attempted to use more energy during the day (73% vs 38%) but less likely to use less energy during the evening (61% vs 89%). However, 61% is still a large fraction given that there was no direct financial benefit for the solar sponge participants to decrease energy use in the evening, and this perhaps explains why solar sponge groups displayed the most obvious changes to their consumption profiles, both increasing daytime consumption and decreasing evening consumption. This contrasts with the peak shave groups, where a large share of responses claimed to have decreased their evening energy use even though no observable response was detected in the actual energy use data.

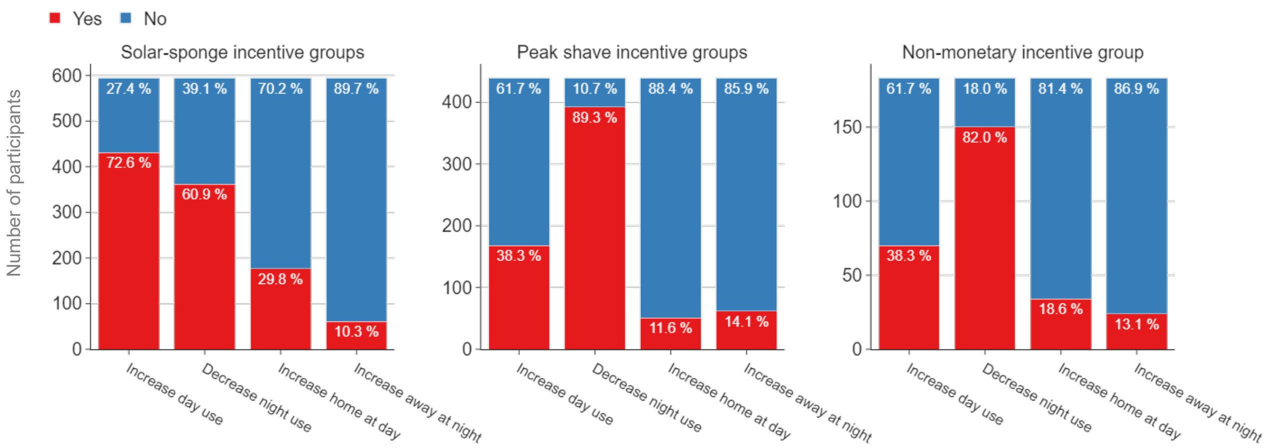


Figure 7-7: Share of participants reporting attempts to increase or decrease the amount of energy they used during the day or night; Non-solar households

[22] Similarly, 95% of solar participants allocated to a peak shave group and 93% of those allocated to the non-monetary group stated they would participate if the program is offered to them again.

[23] The solar users were more similar to each other, with 85% of both peak shave and non-monetary participants stating that they attempted to use more energy during the evening.

Next, motivations to participate in the Load-shifting Challenge were examined. 42% of solar households reported environmental reasons being more important than financial reasons for motivating their load shifting. Only 37% of non-solar households reported the same motivation. This slightly higher propensity for solar households to state an environmental motive for their load-shifting actions aligns with the empirical observation that solar households responded to both monetary and non-monetary peak shave incentives whereas the non-solar households did not.

The survey data also provided insights into how certain household characteristics differed across solar and non-solar households. These characteristics may contribute or suggest mechanisms for observed differences in the magnitude of load shifting between solar and non-solar households. Table 7-4 compares exit survey results across solar and non-solar households. These results suggest that solar households may have a greater ability to respond to load shifting programs given their greater propensity to work from home (37% vs 33%), to use load timers (46% vs 36%), and to use heating and cooling appliances to load shift (66% vs 63%). Solar households might also have slightly higher levels of energy cognizance, reporting more educational attainment (88% vs 84%) and engagement with retail electricity markets (52% vs 46%), but perhaps less financial motivation to load shift as they less frequently report experiencing bill distress. Similarly, non-solar household's higher financial motivations to load shift may also reflect that they were more budget constrained.

Taken together, the self-reported motivations, financial capacity, energy cognizance, and ability to load shift, all suggest that solar households are perhaps more likely to engage in load shifting, particularly where there is no explicit financial incentive to do so, which is consistent with empirical observations.

Table 7-4: Comparison of means for various household characteristics across solar and non-solar households

	Solar	Non-Solar	p-value
Completed survey	0.66	0.65	0.61
Financial			
Experiences energy bill distress	0.17	0.25	0.00
Financial motive to load shift	0.43	0.46	0.09
Energy Cognizance			
Education > HS	0.88	0.84	0.00
Searches retail market each year	0.52	0.46	0.00
Ability to load shift			
Work from home >2 days	0.37	0.33	0.06
Children at home	0.58	0.53	0.01
Used load timers	0.46	0.36	0.00
Used heating/cooling to load shift	0.66	0.63	0.07

Finally, the survey data provided insights into how the use of certain appliance features changed during the Load-shifting Challenge, in particular the use of load timers. 41% of respondents reported using a load timer during the trial. Figure 7-8 describes the appliances used by households with load timers to facilitate load shifting. Most of these households had a time-programmable dishwasher, washing machine or heater/cooler. About 50% reported their dryer having the functionality (which is relevant to this trial period as it covered autumn and a small amount of winter), but significantly less had pool pump or electric water heaters. The red portion of each column reflects the fraction of users that reported that they increased the use of their load timer in the trial compared to before the trial, with 30-40% of these households using the feature more on the appliances where the functionality exists. This shows that a significant share of households used load timers on their appliances and increased their use of these features, demonstrating one vehicle for behaviour change from our trial.

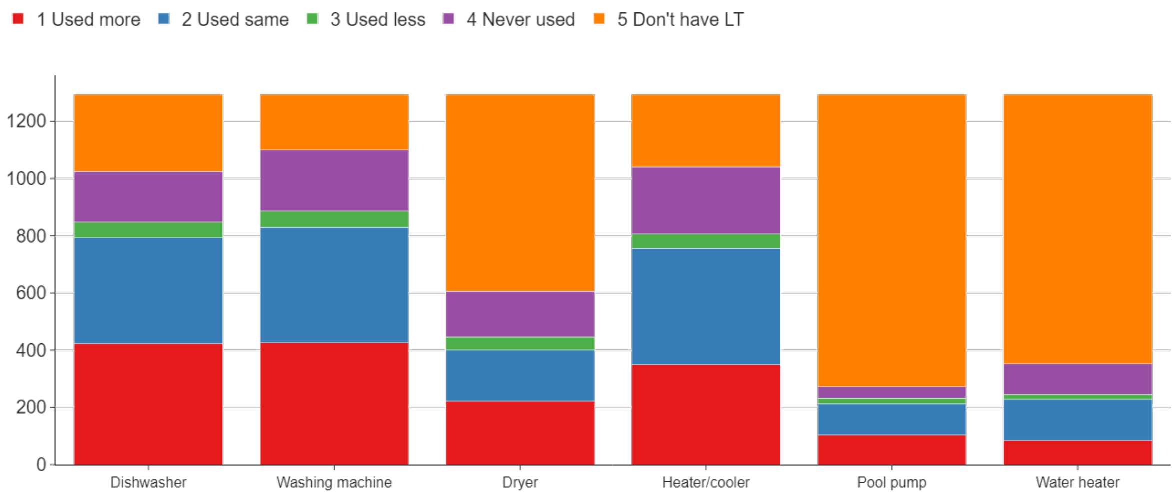


Figure 7-8: Number of participants that used the load timer on each appliance (sample restricted to those that reported having used a load timer during the trial)

8 Recommendations for Next Steps

The Load-shifting Challenge has documented a series of opportunities for industry and policy alongside a raft of further research questions.

Industry and policy

The trial demonstrated how technology providers can leverage the investment made in smart meters to impact the energy-use patterns of households. For example, the simple trial programs that encouraged either more daytime electricity consumption or less evening consumption were successful in changing electricity use patterns for some segments of the population. Retailers can partner with third parties such as Powerpal to implement programs designed to shift energy use or communicate dynamic changes to their tariff rate. Network companies or policy-makers that aim to deliver more efficient price signals and directly incentivise households to load shift can partner with a third party such as Powerpal and side-step partnering with retailers if they are concerned that their network charges or program incentives will not be passed through to households.

It follows that third-party providers, such as Powerpal, can facilitate programs that provide incentives for households to alter their electricity consumption in response to network and wholesale market conditions, perhaps independently of retailers with significant changes to load profile demonstrated with day-ahead event notice. If demand across groups of households is forecastable with reasonable accuracy, then it may be that a third-party can create a service that is eligible for wholesale demand-response programs. This suitability may improve as timed device penetration and understanding improves.

Further research

This research has demonstrated the willingness of some households to engage in load-shifting programs, in some cases in response to financial motives, in other cases in response to non-monetary incentives. It has shown that these responses significantly differ along one characteristic—whether the home has rooftop solar— suggesting that a deeper investigation into how different consumer segments respond to these programs may be warranted. Further, the research was conducted during autumn and part of winter in a summer-peaking grid. Coincidentally, it also overlapped with a period of market difficulties culminating in the June 2022 electricity market suspension, however it is reasonable to expect that these programs could have a greater impact during summer months and there is an opportunity to re-run this trial during those months.

There are also questions regarding the mechanisms that allow for a successful load shifting program. While this work provided some survey evidence that the program used motivated households to more frequently use timer devices on their appliances, direct validation is needed to confirm this. Indeed, arriving at what might be considered an ideal way for household behaviour to change to help accommodate more renewables by automating appliances and having their output vary with market or network conditions is a process requiring many steps. This research demonstrated the willingness of some households to engage with load shifting in an unspecified manner; a logical next step is to investigate their willingness to participate in programs built around device automation (such as pre-heating / pre-cooling homes, water, electric vehicles etc.).

9 Conclusions

This research involved the design, trial, and evaluation of the Load-shifting Challenge, an incentive program aimed at encouraging households to shift their within-day electricity use to align with solar energy output. The results of the randomised control trial demonstrate that attempting to engage households in load shifting may be a valuable exercise for policy-makers aiming to facilitate the integration of renewable energy, or for industry participants trying to create value and capture some surplus presumably to be split with their customers. The value opportunities in Australia from promoting more flexible demand or different usage patterns have steadily increased in recent years. Increased solar penetration, in particular, has resulted in suppressed daytime prices and increased evening prices.

The results showed that simple, and potentially small, incentives can change aggregate consumption profiles among our participants, whether offered every day or on an ad hoc basis. Changed consumption profiles on event days prompted by ad hoc incentives demonstrate the potential to generate flexible demand response.

The observed responses to routine incentives suggest that time-of-use-style tariffs can change load shapes and lower the average cost of supplying households. This suggests that a retailer-customer surplus can be created if customers move from time-invariant fixed rates to time-of-use fixed rates. However, these tariffs do not provide demand response because they do not promote flexibility with respect to real-time conditions. On the other hand, ad hoc events can be called when market or network conditions are such that changing load profiles is forecast to be particularly valuable on specific days. These results demonstrate that meaningful consumption responses can be generated by households on ad hoc event days, and of similar size to responses generated by households that receive every day incentives.

Further, this research clearly documents the importance of targeting and tailoring programs to users. A key characteristic of households in the context of energy pricing programs is whether they have rooftop solar. Participating households without solar installations were shown to be more sensitive to incentive design than solar households. For example, monetary and non-monetary peak shave incentives did not elicit a response for non-solar households for incentives in the range of 0c/kWh to 5c/kWh.

Solar households responded to either form of incentive by decreasing their peak usage by 9%. This suggests that this class of customer is motivated to act by being in a program but without a detectable degree of price elasticity from the 0c/kWh to 5c/kWh incentive range.

Non-solar households did, however, respond to monetary solar sponge incentives where they were paid to use more energy in the middle of the day. Specifically, these households increased energy use into the middle-of-day incentive window by an average of 6%. They also decreased energy consumption during the evening peak by an average of 8%, despite not receiving a direct incentive to do so. These results are encouraging for policy-makers and retailers that aim to reduce the average cost of energy for these customers.

Finally, non-solar households were found to respond to peak shave incentives only when the reward was substantial (50c/kWh). Hence, non-solar households engaged with and responded to some but not all load shifting incentives, emphasising that benefits can be improved from careful incentive design and targeting.

Given that participant consumption profiles were impacted by many of the incentives offered under the Load-shifting Challenge, wholesale price conditions and patterns can be envisaged whereby these incentives will generate economic value and a surplus for both households and retailers. Despite the short trial duration, some treatments were shown to have decreased wholesale procurement costs under prevailing wholesale market conditions, providing a simple demonstration of the potential for these types of programs to create value. It follows that there are opportunities for industry or policy-makers to partner with third-party electricity data providers that directly communicate with households to implement load-shifting programs that create value or achieve policy goals.

References

Australian Energy Market Operator. (2022). 2022 Electricity Statement of Opportunities. A report of the National Electricity Market, released on 31/8/2022.

Australian Photovoltaic Institute (2022) https://pv-map.apvi.org.au/analyses

Allcott, H. (2011). Rethinking real-time electricity pricing. *Resource and Energy Economics, 33*(4), 820-842.

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review, 104*(10), 3003-3037.

Andersen, L. M., Hansen, L. G., Jensen, C. L., & Wolak, F. A. (2021). *Can incentives to increase electricity use reduce the cost of integrating renewable resources*. Retrieved from

Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the new York Academy of Sciences, 1369*(1), 24-39.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*(25), 10367-10371.

Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy, 96*(4), 675-700.

Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2022). *Do the effects of social nudges persist? Theory and evidence from 38 natural field experiments*. Retrieved from

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences, 116*(12), 5293-5298.

Burkhardt, J., Gillingham, K., & Kopalle, P. K. (2019). *Experimental evidence on the effect of information and pricing on residential electricity consumption* (0898-2937). Retrieved from

Bushnell, J., & Novan, K. (2021). Setting with the Sun: The impacts of renewable energy on conventional generation. *Journal of the Association of Environmental and Resource Economists, 8*(4), 759-796.

Byrne, D. P., Goette, L., Martin, L. A., Delahey, L., Jones, A., Miles, A., . . . Tiefenbeck, V. (2021). *The Habit-Forming Effects of Feedback: Evidence From a Large-Scale Field Experiment*.

CME. (2017). *The retail electricity market for households and small businesses in Victoria*.

Costa, F., & Gerard, F. (2021). Hysteresis and the welfare effect of corrective policies: Theory and evidence from an energy-saving program. *Journal of Political Economy, 129*(6), 1705-1743.

Deryugina, T., MacKay, A., & Reif, J. (2020). The long-run dynamics of electricity demand: Evidence from municipal aggregation. *American Economic Journal: Applied Economics, 12*(1), 86-114.

Faruqui, A., & Sergici, S. (2010). Household response to dynamic pricing of electricity: a survey of 15 experiments. *Journal of regulatory Economics, 38*(2), 193-225.

Herter, K. (2007). Residential implementation of critical-peak pricing of electricity. *Energy Policy, 35*(4), 2121-2130.

IRENA (2021), Renewable capacity statistics. (2021). International Renewable Energy Agency (IRENA), Abu Dhabi.

Ito, K., Ida, T., & Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy, 10*(1), 240-267.

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

Jha, A., & Leslie, G. (2021). Start-up Costs and Market Power: Lessons from the Renewable Energy Transition. *Available at SSRN 3603627*.

Kessels, K., Kraan, C., Karg, L., Maggiore, S., Valkering, P., & Laes, E. (2016). Fostering residential demand response through dynamic pricing schemes: A behavioural review of smart grid pilots in Europe. *Sustainability, 8*(9), 929.

Roberts, M., Passey, R., Adams, S., Whittaker, L., Russell-Bennett, R., McAndrew, R., . . . Ben-David, R. (2021). Opportunity Assessment Report: Rewarding flexible demand: Customer friendly cost reflective tariffs and incentives. *RACE for, 2030*.

Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review, 67*(2), 76-90.

Wolak, F. A. (2007). Residential customer response to real-time pricing: The anaheim critical peak pricing experiment.

Wolak, F. A. (2010). An experimental comparison of critical peak and hourly pricing: the PowerCentsDC program. *Department of Economics Stanford University*.

Wolak, F. A. (2011). Do residential customers respond to hourly prices? Evidence from a dynamic pricing experiment. *American Economic Review, 101*(3), 83-87.

Appendix A: General Load-shifting Challenge invitation and welcome email to the non-monetary peak shave group

Load-shifting Challenge invitation

powerpal

Hi *|FNAME|*

We're excited to announce Powerpal has partnered with Monash University to make a new Load Shifter challenge available within the Powerpal app.

Load Shifting is a new way for households to help balance the electricity grid as more low-cost renewable solar and wind energy comes online. By joining in the Load Shifting challenge you'll not only be using energy when it is cleanest but you'll also get the chance to earn real rewards - it's both a win for the planet and for your wallet.

Participating is easy and we'd love for you to join!

Here's how it works:

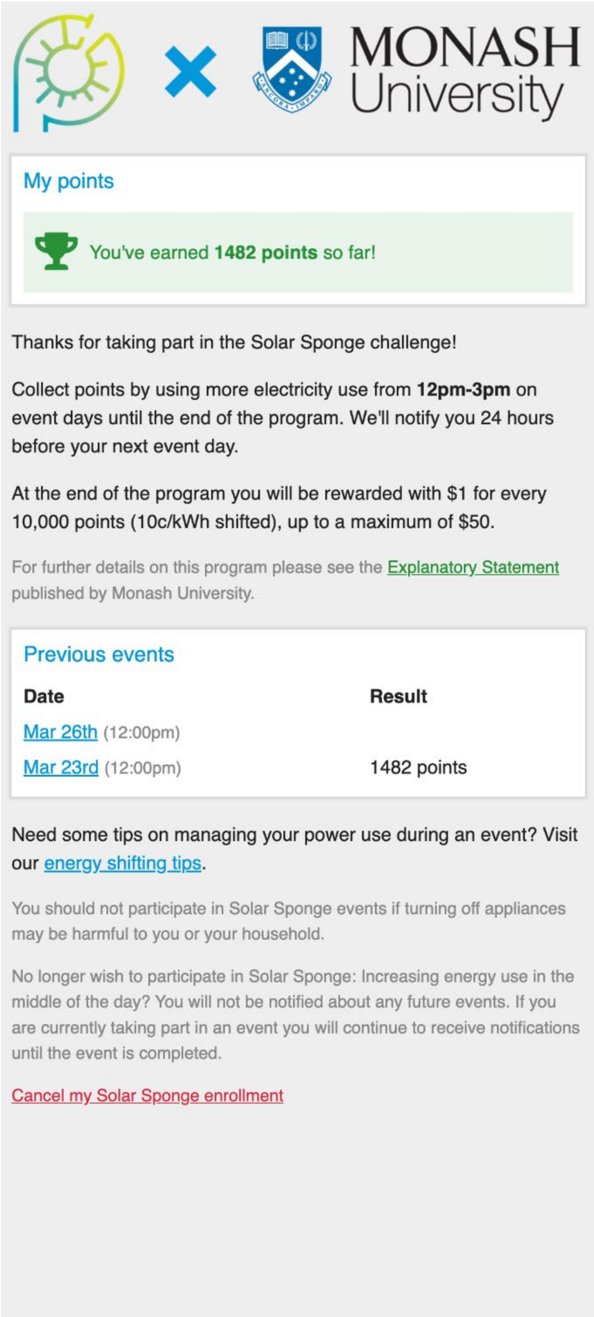
1. **Register**: Join the Load Shifter challenge in the Powerpal app for your chance to start earning rewards.
2. **Get notified**: When there are times of particularly low or high demand on the grid we'll let you know via an alert to your phone so you prepare for the Load Shifting event.
3. **Respond**: Adjust your energy use during the event window for that day - for example, you could turn up or down your thermostat or schedule your next laundry load to match the challenge time window.
4. **Claim your reward**: Sync your energy consumption with your Powerpal app within 48 hours of an event to claim rewards for adjusting your energy consumption. If you're at home during the event you can also track how you're doing in real-time during an event by visiting the challenge dashboard in the Powerpal App.
5. **Tell us how it went**: This challenge will run for around two months. When the challenge finishes we'll ask you to complete a short questionnaire, after which you'll be able to redeem your reward as a Prezzee digital shopping voucher that can be spent at hundreds of major retail stores.

Everyone who registers for the challenge will receive at least a $5 reward and you could even earn up to $50 depending on how much electricity you shift during the Load Shiting events.

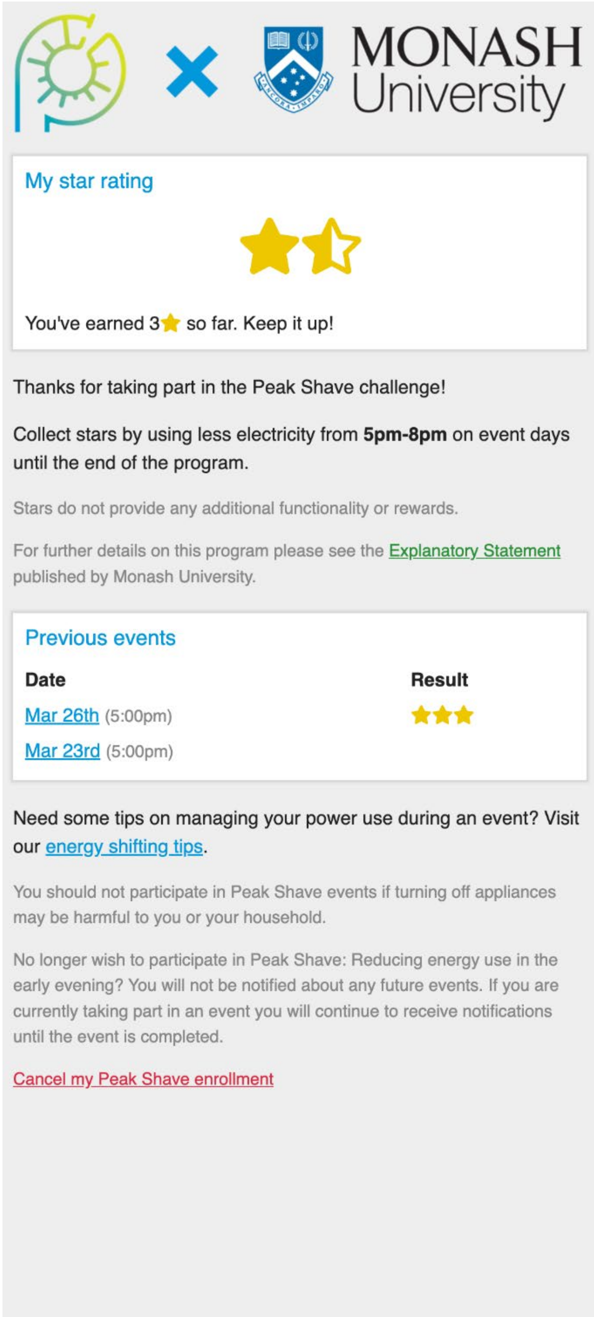
For more information and to register, simply open the Powerpal app and follow the instructions in the Challenges tab. Places are limited so don't delay!

Appendix B: Screenshots of within-app experience for solar sponge and non-monetary groups

Solar sponge group within-app experience



Non monetary group within-app experience



Appendix C: Screenshot of within-app load shifting tips

powerpal

Load Shifting Tips

Obviously, there are some things you can't schedule, like your fridge running or when you cook dinner. However, we've put together some simple tips to help you become a load-shifter.

1. Cooling or heating your house before you get home.

Use your device's inbuilt timer to heat or cool your house before you get home. Heating and cooling make up to 40% of your household energy bills so shifting this usage can pay huge dividends over time.

2. Use timers or delay functions to shift optional, non-time-sensitive use to the middle of the day.

For example:

- **Put a timer on your dishwasher** and let it run during the middle of the day.
- **Load up your washing machine** and use a timer to run it during the day. If you're using a dryer, use it during the weekend when you're home during the day.
- If you own a pool, time your **pool pump** to run during times of peak solar.
- Charge your **Electric Vehicle** in the middle of the day, not during the evening.

3. Practice conservation during peak times

During peak times, it's important to keep on top of your energy use. Make sure to turn off your lights, TVs and electronics when they're not being used. Consider delaying your use of appliances, or if you have to use them make sure to use their eco-modes. For more information on the main sources of energy use in your household check out our article.

Using an electricity usage meter like Powerpal can provide real-time feedback on your usage.

Conclusion

Load-shifting isn't complex, and the small and easy to execute changes above can help you load-shift and ultimately create a more stable and renewable friendly energy grid for all of us. Real-time feedback is important to this process and energy monitors like Powerpal are invaluable in taking control of your energy usage.

RACE for 2030

www.racefor2030.com.au

Australian Government
Department of Industry, Science, Energy and Resources

AusIndustry
Cooperative Research Centres Program