Pick One or Both - Contrasting and Combining HPWH load Control Approaches

Stephan Barsun, Jenny Hughes, Verdant Associates
Amélie Besson, Association for Energy Affordability
Emily Kehmeier, Daniela Urigwe, Energy Solutions

ABSTRACT

Heat pump water heaters (HPWHs) can shift electric load from high-demand periods to lower-demand hours. Realizing this potential, however, depends on load-shifting strategies that are both reliable and consistently implemented.

This paper compares two HPWH load-shifting approaches: scheduled Time-of-Use (TOU) control under the Self-Generation Incentive Program HPWH effort (SGIP HPWH) and signal-based control in Pacific Gas and Electric's WatterSaver pilot. Both approaches seek to move consumption away from peak hours, but they differ in implementation, flexibility, and operational risk.

Under the SGIP HPWH program, contractors program TOU schedules using the manufacturers' protocols into units at installation, enabling autonomous peak-period demand reduction without internet connectivity. WatterSaver instead uses daily remote commands to initiate either Basic Load Up (BLU) or Advanced Load Up (ALU), with ALU temporarily increasing tank temperature to support deeper peak-period load shedding.

Both approaches reduced peak-period consumption. SGIP devices with verified settings achieved average reductions of 0.37 kWh during the 4–9 PM weekday peak, compared with 0.25 kWh for WatterSaver BLU and 0.47 kWh for WatterSaver ALU.

However, implementation failures limited realized performance. Misconfiguration affected 27–44% of sampled SGIP devices from one manufacturer due to a contractor app bug, while 15–32% of WatterSaver commands failed during the analysis period. These findings show that reliable setup, verification, and communication are essential to realizing HPWH load flexibility benefits. Scheduled control may offer a lower-cost and simpler pathway but may provide less flexibility and smaller peak impacts than signaled approaches.

Introduction

This paper presents the impact from two overlapping and complimentary heat pump water heater (HPWH) load shifting programs:

- The Self-Generation Incentive Program (SGIP HPWH) program that uses ***scheduled*** TOU load shifting
- Pacific Gas and Electric's (PG&E's) WatterSaver program that sends daily ***signals*** to participating HPWHs

SGIP HPWH (Scheduled Load Shifting) Program Overview

The SGIP HPWH program incentivized the installation of over twelve thousand HPWHs from 2023 through the end of 2025. These were divided across the single family, multifamily and commercial sectors. These HPWHs are further divided by type (unitary vs. central) and if they are serving general or equity markets. Unitary HPWHs are like traditional tank water heater in many homes with a single tank and the heat pump and other heating elements connected to that tank at the factory and installed as a single unit by the installer. These serve single family homes or (usually) a single unit in multifamily residences. For this paper, we will only present findings for units serving single family residences.

All SGIP HPWH program participants were required to integrate demand response and load shifting. For this program, Joint Appendix 13 (JA13) of California's Title 24 was selected as the guideline to govern load shifting for unitary systems.

JA13 specifies a default TOU schedule that water heaters must follow, defining the hours during which load shifting must occur (generally the utility peak window, 4-9PM in PG&E territory). JA13 is configured by the HPWH contractor using the manufacturer's setup app at the time of installation, and the schedule is stored in the water heater's firmware. Once set, the HPWH follows this schedule autonomously without any internet connection or ongoing communication—it simply knows what time it is and acts accordingly. Manufacturers are permitted to develop their own control algorithm within the JA13 framework, meaning the precise logic for how the unit reduces consumption during shed window can vary by brand. JA13 schedules can, in principle, be modified or overridden by a user (for a 72-hour period) or by the manufacturer (if there is an internet connection to the unit, among other limitations). There is currently no standard to confirm JA13 status for an enrolled site, nor is there 3rd party validation of JA13 manufacturer settings. More information about the implementation of the SGIP HPWH is available in Kehmeier et al, 2026.

WatterSaver (Signaled Load Shifting) Program Overview

Launched in 2022, with the first water heater enrolled on March 20th, 2022, the WatterSaver Pilot Program manages electric water heaters with the goals of shifting energy usage outside peak periods, enhancing grid reliability, and reducing costs for PG&E's Time-of-Use (TOU) customers. PG&E contracted with the Association for Energy Affordability (AEA) to conduct the implementation of the WatterSaver Pilot. Virtual Peaker (VP) is the third party Distributed Energy Resource Management System (DERMS) provider that actively controls enrolled participants. These controls consist of two types of load shifting commands sent daily to participating water heaters:

- Load Up – the water heater is sent a command to heat the stored water 1) to the customer's desired temperature set point by reducing the deadband before the shed period told to store more energy either by bringing more of the tank to the set point temperature (Basic Load Up or BLU) or to 2) to a tank temperature that is higher than the customer's set point (Advanced Load Up or ALU).
- Shed—The water heater is sent a command to reduce energy consumption; each manufacturer uses its own logic to determine how this is done, but it usually involves lowering the setpoint, lowering the bottom bound of the deadband that the tank

temperature can drift to, and/or using more efficient operating modes. The shed commands are the same for both BLU and ALU. Shed periods usually align with the participant's peak tariff hours.

In terms of control behavior, BLU closely parallels the SGIP HPWH's JA13 scheduled approach: primarily relying on the stored thermal mass to coast through peak hours. The primary difference is that JA13 operates on a fixed autonomous schedule, while BLU is triggered by a daily signal and incorporates a small degree of 'loading up' more of the water to the water heater setpoint. ALU goes further by raising the tank temperature above the customer's setpoint before the shed window, storing more thermal energy and enabling a longer or deeper shed. ALU is best suited for participants on rates with a large peak-to-off price differential, where the energy cost savings during the shed outweigh the additional energy consumed to super-heat the tank (Besson, 2024). ALU can mitigate snapback but is only available to customers with a thermostatic mixing valve, since in absence of one there is an increased scalding risk. BLU is appropriate for participants on standard TOU rates or where tank capacity or safety constraints limit how high the setpoint can be raised.

Signaled control introduces points of failure that embedded JA13 scheduling does not: the implementer must generate a command, the command must be successfully transmitted to the device, and the device must accept and execute the commands. Additionally, the end user must not opt out of the commands. Failed commands, whether due to communication errors, device errors, scheduled vacation mode, or participant opt outs, result in the water heater operating on its default schedule rather than the signaled schedule, reducing the program's load shifting impact. Since command failure can happen for multiple reasons, we treat days with a failure detected separately from both baseline days and days where commands were executed successfully. Note that more information about the WatterSaver program may be found in Besson, 2024 and Besson, 2026.

Methodology

SGIP HPWH Impact Methodology

To estimate the impact of scheduled load shifting on the timing of SGIP HPWH energy usage, the team initially aimed to recruit and install CTA-2045[1] modules on a sample of enrolled devices and implement a randomized control trial (RCT) to shut off JA13 for sites that were randomly selected each day, creating a baseline for comparison to. We determined that this would not work as we were not able to control load shifting in this method, so we pivoted to Plan B, working with manufacturers to obtain data and have them implement shut off periods where a baseline for each site would be created.

We used telemetry data received from one of the more prolific SGIP HPWH manufacturers as part of our 'Plan B' data collection, with data for this paper from January 1, 2025 to April 30, 2025. Under Plan B, the manufacturer temporarily paused JA13 schedules on the HPWHs for customers who agreed to participate in our study. The manufacturer provided Verdant with telemetry data for these customers for periods prior to the pause, during the pause,

[1] CTA-2045 is a modular communications interface standard that enables two-way communication between utilities and smart appliances for energy management and demand response.

and following the pause in the JA13 schedule. The paused period provides the evaluation with baseline energy usage when the HPWH is not operating under the SGIP prescribed JA13 schedule, while the pre- and post-pause periods of data represent data impacted by the SGIP program. Using the baseline and impact period data, it is possible to estimate the load shifting impacts of the program.

To estimate the load shifting impact, we used hourly panel data regression models with fixed effects. This approach uses data with both a time series (multiple time intervals on a given water heater) and cross sectional (multiple water heaters) dimension to model shifts in water heater energy consumption due to JA13 load shifting. The model includes a binary variable to identify the baseline versus the program periods and the estimated coefficients determine the hourly impacts of the JA13 load shifting settings on the water heating energy usage.

As described above, by implementing the "baseline period" pause in the JA13 schedule, the manufacturer and evaluation team determined that many of the SGIP HPWHs were not following the JA13 schedule. For the impact analysis, we estimate the impact of load shifting on the sample of 50 water heaters who agreed to participate in our study and separately for the smaller sample of water heaters where the JA13 schedule was verified. The methodology for estimating these impacts is described in the Joint Methodology section below.

WatterSaver Energy Impacts Methodology

To model WatterSaver impacts, we used 15-minute interval telemetry and daily event log data from May 1, 2023, to April 30, 2024. These data were reviewed, and quality controlled for use in the analysis. All WatterSaver participants are automatically enrolled in an RCT, where each day a random 10 percent of sites in a given platoon (such as BLU 4-9PM) are not sent load shifting commands, creating a baseline.

To estimate WatterSaver program ex-post impacts, Verdant utilized hourly panel data regression models with participant fixed effects. This approach uses data with both time series (datetime) and cross sectional (household) dimensions to model water heater consumption. Participant-fixed effects are included in the model to account for variability by household. Participants are segmented into modeling groups based on platoon and hour of the day. Independent variables in the models include weekday/weekend, and device communication status (if the command message was successful or if it failed). The results of the panel models are used to determine if WatterSaver water heaters were successful at moving energy usage away from the shed hours, and to determine if the water heaters use more energy on days when they load shift versus days where they did not load shift. These results ultimately help identify if load up and shed strategies should be adjusted to optimize efficiency and load shifting impact. Due to data availability when this analysis was performed, we could not account for seasonality or model by any additional segmentation within the platoon for variables such as climate zone or water heater manufacturer. The model specification is presented in the Joint Impacts Methodology section below.

Joint Impacts Methodology

Both the SGIP HPWH and WatterSaver program evaluations use the same core modeling framework to estimate load shifting impacts: hourly fixed effects panel data regression models estimated separately for each hour of the day, producing 24 hourly coefficient estimates of load shifting impact in kWh. This approach uses data with both a time series dimension (multiple observations per water heater over time) and a cross-sectional dimension (multiple water heaters) to model shifts in water heater energy consumption. Modeling each hour separately enables the models to capture independent time-of-day impacts, including load up (for ALU WatterSaver devices only), the shed during peak hours and the subsequent snapback as tanks reheat. For each hourly regression, observations are indexed by device I and day t, where a day is defined as 10am – 10am the following calendar day so that any post-peak snapback fails within the same daily unit as the corresponding load.

The approach for each project controls for individual device fixed effects, which capture time-invariant characteristics specific to each water heater—such as household size, occupant water use habits, and equipment efficiency—that would otherwise confound impact estimates. Both models also include weekday/weekend indicator variables to account for systematic differences in usage patterns.

The two approaches differ in three respects, driven by differences in how each program is designed and how the baseline is established.

1. The model used to evaluate WatterSaver impacts includes daily fixed effects (Γ_T), which account for variation common across all water heaters on a given day, such as weather conditions. This is possible because WatterSaver operates as an ongoing program with a RCT design, where baseline days occur continuously alongside controlled days, allowing day-level variation for shift/not shift to be isolated. The SGIP HPWH evaluation cannot include daily fixed effects because the baseline was established through a discrete pause period during which all study devices were simultaneously set to baseline status; meaning there are no concurrent baseline days happening while other sites are shifting. To partially account for weather in the SGIP HPWH model, an ambient temperature control is included, capturing the effect of room temperature at the HPWH install location.
2. Because WatterSaver commands can fail, meaning that at a given site on a given day the HPWH is shifting successfully, baselined, or failing to shift, the WatterSaver model includes a separate term for failed command days, distinguishing the impact of a successful control event from a failed one. The SGIP HPWH program has no equivalent, these devices are either properly configured at setup or were not properly configured so essentially should show no program impact.[2] The failure to configure devices will be accounted for in the impacts by showing the impact for the entire sample and for just the devices that were configured to shift with JA13.
3. WatterSaver participants are divided into multiple platoons, BLU 4-9PM (n=74), ALU 4-9PM (n=86), and ALU 3-12AM (n=21) and separate models are estimated for each platoon. The SGIP HPWH analysis covers a single manufacturer and TOU schedule type, analogous to a single platoon (n=50).

The model specifications for SGIP HPWH and WatterSaver are presented in Equations 1 and 2, respectively, below.

[2] The reconfiguration of devices with known issues occurred after the analysis period for this paper. Corrections for HPWHs connected through the manufacturer app were implemented via a remote update. For devices without an app connection, the extent of incorrect configurations is unknown, and remote reconfiguration was not possible.

Equation 1: Model Specification for SGIP HPWH Analysis

$$kWh_{it} = \alpha_i + \beta_S Shift_{it} + \beta_W Weekend_{it} + \beta_{SW}(Shift * Weekend)_{it} + \beta_{AT}(Ambient\ Temp)_{it} + \varepsilon_{it}$$

Where:
kWh_{it}= Actual water heater kWh usage for customer *i* during date interval *t*
Shift = Indicator variable (0, 1) equal to 1 on shift days and 0 on baseline days
Weekend = Indicator variable (0, 1), where 1 is a weekend date
$\beta_S,\ \beta_W,\ \beta_{SW}$ = Coefficients estimated by the model, representing the effect of shifted days, weekend status, shifted day and weekend status interacted, respectively
α_i= Water Heater-specific fixed-effect, capturing all time-invariant characteristics unique to the customer
β_{AT} = Effect of ambient temperature, capturing the impact of the ambient room temperature where the HPWH is located
ε_{it}= Random error term, assumed to be normally distributed

Equation 2: Model Specification WatterSaver

$$kWh_{i,t} = \alpha_i + \Gamma_T + \beta_C Success_{i,t} + \beta_F Fail_{i,t} + \beta_W Weekend_t + \beta_{CW}(Success * Weekend)_{i,t} + \beta_{FW}(Fail * Weekend)_{i,t} + \varepsilon_{i,t}$$

Where:
$kWh_{i,t}$= Actual water heater kWh usage for customer *i* during date interval *t*
Success = Indicator variable (0, 1) for command success
Fail = Indicator variable (0, 1) for command fail
Weekend = Indicator variable (0, 1), where 1 is a weekend date
$\beta_C,\ \beta_F,\ \beta_W\ ,\beta_{CW},\ \beta_{FW}$ = Coefficients estimated by the model, representing the effect of successful events, failed command events, weekend status, successful command and weekend status interacted, and failed command and weekend status interacted, respectively
α_i= Water Heater-specific fixed-effect, capturing all time-invariant characteristics unique to the customer
Γ_T = Daily fixed effect, capturing any variation that is common across water heaters but vary over time
$\varepsilon_{i,t}$= Random error term, assumed to be normally distributed

For the SGIP HPWH model, all impacts are estimated relative to baseline usage on a weekday. The coefficient β_S estimates the kWh impact of a shifted day on weekday water heater usage. Weekend baseline usage is estimated as β_W, and the weekend shift impact relative to that weekend baseline equals $\beta_S + \beta_{SW}$.

For the WatterSaver model, impacts are similarly estimated relative to weekday baseline. The coefficient β_C estimates the kWh impact of a successful command day on weekdays. The weekend impact of a successful command is $\beta_C + \beta_{CW}$, relative to weekend baseline. The weekday impact of a failed command day is estimated by β_F, and the weekend impact is $\beta_F + \beta_{FW}$.

Neither model accounts for seasonality. The SGIP HPWH baseline period covers winter and spring months only, so seasonal extrapolation is not appropriate at this stage. The final evaluation report will incorporate additional baseline periods to enable seasonal analysis. WatterSaver results reflect approximately one year of data and are less sensitive to this limitation, though climate zone and manufacturer segmentation were not available for this analysis.

Results

Both the SGIP HPWH and WatterSaver programs demonstrate meaningful load shifting during peak periods, with statistically significant reductions in energy use during the shed periods followed by a snapback as tanks reheat. A consistent finding across both programs is that device-level failures—whether JA13 schedules that were never properly configured (SGIP HPWH) or unsuccessful command delivery (WatterSaver)—reduce realized impacts relative to what fully functioning devices achieve. When devices are confirmed to be operating correctly, the SGIP HPWH scheduled approach and WatterSaver's BLU signaled approach produce comparable weekday shed margins, as shown in the head-to-head comparison in Table 4. The sections below present results for each program before turning to that direct comparison.

SGIP HPWP (Scheduled) Impacts

A significant finding emerged during data collection: only 56% of study sample HPWHs were confirmed to be following a JA13 schedule. Because of this, we present results in two ways: for the full sample of 50 sites, and for the 28 confirmed JA13 enabled sites. A review of all cloud-connected heat pump water heaters (HPWHs) by this manufacturer—about one-quarter of installed units—found that 73% were correctly implementing JA13 load-shifting schedules, compared with 56% in the evaluation study sample. These lower-than-expected implementation rates were traced to a flaw in the contractor setup app, which often appeared to display a successful JA13/TOU enrollment without fully enabling the schedule. Although the program required screenshot verification from the contractor for the application, this process did not reliably confirm schedules were implemented given that display issue. The evaluation team worked with the manufacturer to better understand this issue, and the manufacturer deployed an update to correct JA13 setup for cloud-connected units and fixed the contractor app. This occurred after the study period included in the paper. It is possible that requiring more rigorous JA13 testing would have identified this issue much earlier. However, a post-update review indicated that some HPWHs continued operating on custom user schedules, which can override JA13 settings.

Table 1 shows the modeled impacts for the SGIP HPWH program for the 28 sites where JA13 was successfully enabled and the entire sample of 50 sites that included 22 HPWHs that did not have JA13 enabled. In addition to the kWh modeled impacts, the table also displays the percent of that modeled impact relative to the baseline usage during the period. For example, a percentage of 100 would indicate that all possible load during the period was shifted.

Table 1: Modeled SGIP HPWH Scheduled Shift Energy Impacts

	Day Type	Impact During Specific Periods of the Day (kWh and %tage of baseline)		
		Shed Hours (4-9pm)	Snapback Hours (9pm-2am)	Daily Impact 10am-10am
JA13 Enabled (28)	Weekday	-0.37† (-48%)	0.37† (48%)	0.06 (2%)
	Weekend	-0.36† (-48%)	0.31† (41%)	-0.18† (-5%)
All Sites (50)	Weekday	-0.24† (-23%)	0.22† (24%)	0.03 (1%)
	Weekend	-0.21† (-23%)	0.21† (26%)	-0.02 (-1%)

† indicates significance with a 90% confidence interval
§ Due to rounding, the sum of individual components may not exactly equal the daily total.

Both groups, all 50 sites and the 28 confirmed to have JA13 enabled, show clear and statistically significant load shifting behavior (Table 1). Energy use drops meaningfully during the 4-9PM peak shed window, followed by a snapback in the post load shifting period as the tank reheats. Importantly, the daily net energy impact is near zero on weekdays, indicating that JA13 is successfully moving energy consumption without substantially increasing overall energy consumption.

Sites with confirmed JA13 activation show roughly 50 percent greater per-unit shed impacts than the full sample (on weekdays, a reduction of 0.24 kWh across all sites and 0.37 kWh for sites with JA13 enabled), which is to be expected as we do not expect load shifting impacts from those units where JA13 was not successfully enabled. During shed hours, 48 percent of load was modeled to be shifted relative to baseline usage on both weekdays and weekends for sites with JA13 enabled. The baseline data covers only February-March, so seasonal variation is not yet captured, and results should not be extrapolated to an annual impact. In addition, the sample is not fully representative of the broader SGIP HPWH population; PG&E participants and San Francisco Bay Area climate zones are overrepresented compared to the statewide program. Bill and greenhouse gas impacts have not been estimated at this stage given those limitations. These gaps should be addressed in the final evaluation report.

WatterSaver (Signaled) Impacts

For WatterSaver (signaled) impacts, before examining load shapes, it is worth noting command success rates, which affect how much of the enrolled population is shifting on any given day. Across all platoons, 10 percent of days are baselined (no command is sent), approximately 72-75 percent of daily commands were successfully received and executed by devices, and the remaining 15-18 percent of days the devices failed to shift successfully.

Table 3 presents the daily WatterSaver impacts by time period, platoon, and day type on days where the control command was successful. Similar to Table 1, both the impacts and the percentage that impact is relative to baseline are presented.

Table 2: Modeled WatterSaver Daily Impacts for Successfully Controlled Devices by Segment and Day Type

Segment	Day Type	Impact During Specific Periods of the Day (kWh and %tage of baseline))			
		Load Up Hours	**Shed Hours**	**Snapback Hours**	**Daily Impact 10am-10am§**
HP|BLU|4-9PM*	Weekday	0.05†	-0.25† (-36%)	0.20†	-0.02 (-1%)
HP|BLU|4-9PM*	Weekend	-0.01	-0.55† (-67%)	0.25†	-0.43† (-13%)
HP|ALU|4-9PM**	Weekday	0.46†	-0.47† (-61%)	0.11†	0.13† (4%)
HP|ALU|4-9PM**	Weekend	0.43†	-0.53† (-64%)	0.10†	0.07 (2%)
HP|ALU|3PM-12AM***	Weekday	0.59†	-0.32† (-34%)	0.25†	0.57† (21%)
HP|ALU|3PM-12AM***	Weekend	0.51†	-0.15† (-16%)	0.13†	0.42† (14%)
Weighted Average	**All**	**0.25**	**-0.39**	**0.16**	**0.03**

† indicates significance with a 90% confidence interval
§ Due to rounding, the sum of individual components may not exactly equal the daily total.
*BLU 4-9PM periods are: 2pm-4pm Load Up, 4pm-9pm Shed, 9pm-10pm Snapback, 10pm-2pm Other
**ALU 4-9PM periods are: 1pm-4pm Load Up, 4pm-9pm Shed, 9pm-10pm Snapback, 10pm-1pm Other
***ALU 3PM-12AM periods are: noon-3pm Load Up, 3pm-12am Shed, 12am-1am Snapback, 1am-noon Other

The modeled impacts find statistically significant reductions in usage for all three platoons during the shed hours. The largest reduction in usage is 0.55 kWh for the BLU 4-9PM during the weekends, closely followed by a reduction of 0.53 kWh on weekends and 0.47 kWh on weekdays for the ALU 4-9PM platoon. If the shed is constant over the five-hour window, these sheds are approximately 0.098-0.11 kW reduction per hour per device. Smaller, but statistically significant reductions are found for the BLU 4-9PM platoon on weekdays (0.25 kWh or 0.05 kW/hour) and the ALU 3PM-12AM platoon (0.32 kWh or 0.035 kW/hour weekday and 0.15 kWh or 0.016 kW/hour weekend). However, the kW reductions are not consistent over the shed period, with peak reductions occurring at different times, sometimes occurring at the start of the shed period but often influenced by peak demand in the underlying baseline usage.

The longer shed platoon (ALU 3PM-12AM) has slightly more load up but less reduction in energy during the shed hours than the ALU 4-9PM platoon. The two ALU platoons have statistically significant increases in usage during the load up hours and all three platoons have statistically significant snapback or increase in usage following the shed event. However, this behavior can vary depending on user setpoint (Besson, 2026). Combining the energy impacts for all daily time periods, successfully controlled devices lead to a statistically significant reduction in energy usage for devices in the BLU 4-9PM platoon on weekends. Successfully controlled devices in the ALU 4-9PM weekday and ALU 3PM-12AM weekdays and weekends result in daily increases in energy usage. The impact on energy bills of the daily increase in usage combined with the reduction in usage during peak or shed hours will be a function of impacts and TOU price differentials. Similarly, the impact on emissions of the daily increase in usage combined with the reduction in usage during peak or shed hours will be a function of emissions during those hours.

Table 2 above only shows the impacts of shifting in a best-case scenario where all devices are successfully shifting everyday, which is not what happened in the program. Table 3 shows the daily weekday and weekend kWh modeled impacts by platoon, including the percentage that impact is relative to baseline usage. The results are the daily kWh impacts accounting for the percentage of devices in baseline, controlled, and failed status over the last year of the analysis period.

Table 3: Modeled Daily kWh Impacts by Control Strategy per Device (kWh) Including All Days

Segment	**Day Type**	**Impact During Specific Periods of the Day (kWh and %tage of baseline))**			
		Load Up Hours	**Shed Hours**	**Snapback Hours**	**Daily Impact 10am-10am**
HP|BLU|4-9PM*	Weekday	0.04	-0.20 (-29%)	0.15	-0.04 (-1%)
HP|BLU|4-9PM*	Weekend	-0.02	-0.45 (-55%)	0.19	-0.38 (-12%)
HP|ALU|4-9PM**	Weekday	0.35	-0.38 (-50%)	0.08	0.05 (1%)
HP|ALU|4-9PM**	Weekend	0.34	-0.43 (-52%)	0.06	0.00(0%)
HP|ALU|3PM-12AM***	Weekday	0.45	-0.28 (-30%)	0.17	0.34 (12%)
HP|ALU|3PM-12AM***	Weekend	0.36	-0.11 (-12%)	0.07	0.21 (7%)
Weighted Average	**All**	**0.19**	**-0.32**	**0.12**	**-0.03**

*BLU 4-9PM periods are: 2pm-4pm Load Up, 4pm-9pm Shed, 9pm-10pm Snapback, 10pm-2pm Other
**ALU 4-9PM periods are: 1pm-4pm Load Up, 4pm-9pm Shed, 9pm-10pm Snapback, 10pm-1pm Other
***ALU 3PM-12AM periods are: noon-3pm Load Up, 3pm-12am Shed, 12am-1am Snapback, 1am-noon Other

Advanced Load Up (ALU) for the 4-9PM platoon has more load reduction during the weekday shed period and moves load to earlier in the day (Load Up Hours) as the ALU HPWHs are storing more energy to prepare ahead of the peak period than BLU HPWHs. Also, for the 4-9PM platoons, BLU has more snapback than ALU as these HPWHs need to make up for the energy lost during the shed period since they did not pre-heat like ALU HPWHs.

Comparing ALU 4-9PM vs. 3PM-12AM platoons also presents some differences. The longer shed platoon has slightly more load up but less reduction in energy during the shed hours. This lower reduction is likely due to three reasons:

- The water heater can't 'coast' as effectively over 9 hours vs. 5 hours without turning on;
- The 3PM-12AM platoon consumes less energy during shed hours than the 4-9PM platoon, resulting in less potential to shed load (possibly due to no winter data when water heating loads tend to be higher);
- The 3PM-12AM ALU platoon has a higher command failure rate with a higher fraction due to opt outs (32% vs 16%) than the 4-9PM platoon, reducing shed impacts. Participants in this platoon are on either E-Elec or EV2-A with partial peak periods extending the 4-9PM peak to 3PM-12AM.

Table 4 presents the estimated average yearly energy and bill cost impacts per device by platoon, as well as a weighted average across the 3 platoons.[3]

[3] All results presented in Table 4 were weighted by the average participant counts over the year leading up to this report for all command statuses. Given this aggregated nature, we did not estimate significance but expect the annual changes may not be statistically significant.

Table 4: WatterSaver Impact Results Accounting for Control Success, Failure, and Baseline Days

Segment	Average Yearly Impact per Device		
	kWh Energy	**% of Annual Baseline Energy Impact**	**Bill Impact ($)**
HP|BLU|4-9PM	-50.1	4% reduction	-$24.81
HP|ALU|4-9PM	13.0	1% increase	$13.84
HP|ALU|3PM-12AM	110.5	11% increase	$49.24
Weighted Average	**-11.0**	**1% reduction**	**-$2.22**

BLU HPWHs provide a minor net reduction in kWh energy consumption as well as both customer bill costs. ALU HPWHs have higher overall kWh energy consumption and customer bill costs as the water heaters use more energy across an average day. However, different rate structures with a higher differential between peak and off-peak for example, could reverse the bill increase. ***Note these results combine multiple groups for each segment and may not be statistically significant***. However, the ALU algorithm was refined in December 2024 to make load up more relative and minimize snapback in response to observations of increased bills. The impact of these algorithmic changes is expected to alter bill impacts of ALU moving forward.

Scheduled vs. Signaled Load Shifting

JA13 provides a scheduled TOU control (***scheduled)*** load-shifting approach in which the heat pump water heater (HPWH) automatically follows a TOU-optimized schedule. Because this control logic is embedded within the equipment, JA13 programming does not require internet connectivity or external communication. Each manufacturer is currently free to develop their own algorithm and approach. In contrast, PG&E's WatterSaver and similar Water Heater Rewards programs employ ***signaled*** load-shifting strategies. These programs issue control commands, or signals, from a system such as a DERMS, adjusting water heater operation in near real time. Signaled strategies often include a pre-peak "load-up" period—either Basic or Advanced.[4]

The results displayed in Table 5 are for weekdays only and should be used to contrast which 4-9PM shifting strategy might have the potential to create the most savings. During the shed time of 4-9PM, the ALU WatterSaver devices can generate the largest load-shifting impact at a reduction of 0.47 kWh, which is 61 percent of baseline usage that period. Notably, it is intuitive that this platoon would generate the largest load shifting impact because the HPWH is heated above the user set point in the hours leading to the shed period. The goal of ALU is not only to have the most shed reduction but to also dampen the snapback impact since the water heater is not as likely to need to recover as much energy to return to the user set point because of the advanced load up. There is a drawback for ALU however, in that the daily impact of load shifting is an overall increase in energy usage, meaning that depending on the TOU price differential between peak and off peak hours there is a chance that customer bills might increase,

[4] Basic load up does not increase the water heater set point during the load up period before the shed period. Advanced load up increases the water heater set point up to the maximum allowed by the manufacturer during the load up period, increasing the energy stored in the water heater.

though this is more than offset by the monthly $5 per month program incentive. For BLU and SGIP HPWH, the daily impact of energy usage is not significant.

Table 5: Modeled Daily Weekday Impacts Across Approaches[5]

Type of Control	Program and Control Strategy	Impact During Specific Periods of the Day (kWh and %tage of baseline)			
		Load Up Hours	Shed Hours	Snapback Hours	Daily Impact 10am-10am
Signaled	**WatterSaver BLU 4-9 ***	0.05†	-0.25† (-36%)	0.20†	-0.02 (-1%)
Scheduled	**SGIP HPWH 4-9 ****	N/A	-0.37† (-48%)	0.37†	0.06 (2%)
Signaled	**WatterSaver ALU 4-9 ****	0.46†	-0.47† (-61%)	0.11†	0.13† (4%)

† indicates significance with a 90% confidence interval

§ Due to rounding, the sum of individual components may not exactly equal the daily total.

*WatterSaver BLU 4-9PM periods are: 2pm-4pm Load Up, 4pm-9pm Shed, 9pm-10pm Snapback, 10pm-2pm Other

**SGIP HPWH periods are: 4pm-9pm Shed, 9pm-2am Snapback[6], 2am-4pm Other

***WatterSaver ALU 4-9PM periods are: 1pm-4pm Load Up, 4pm-9pm Shed, 9pm-10pm Snapback, 10pm-1pm Other

Figure 1 presents the average weekday load shapes for SGIP HPWH baseline and shifted days, alongside the corresponding WatterSaver BLU baseline and shifted weekday load shapes. The data for WatterSaver is drawn from Verdant's WatterSaver Interim Evaluation and did not include any SGIP HPWH participants.[7] The shaded region in each figure denotes the 4–9 PM shed window. For WatterSaver, only successful shift events are included, while the SGIP HPWH shifted profile reflects days (and devices) where JA13 was enabled and active. The SGIP HPWH baseline is based on the February-March 2024 period when JA13 was paused. The WatterSaver baseline reflects baseline days from the WatterSaver RCT.

[5] These are still preliminary comparisons that will be expanded and refined in the final WatterSaver report later in 2026. Additionally, scheduled load control may evolve to include ALU with one manufacturer already implementing this in JA13 logic for their newest HPWHs.

[6] SGIP HPWH snapback periods were set longer than WatterSaver based on a longer period with statistically significant differences to baseline.

[7] Verdant plans to complete the final WatterSaver Evaluation Report later in 2026.

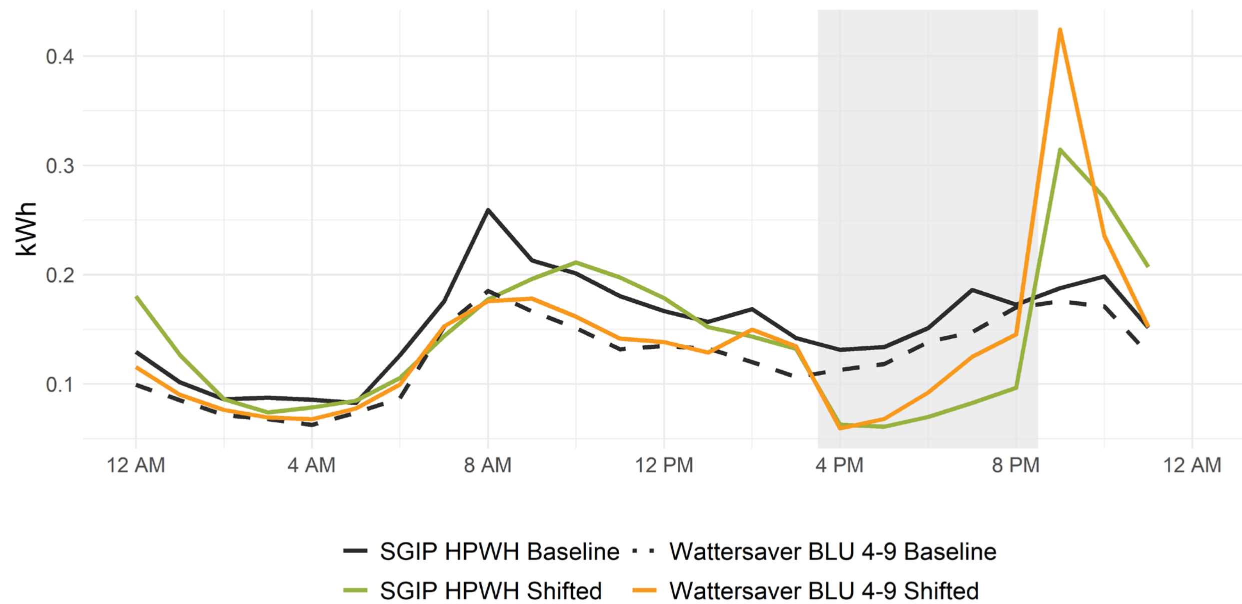

Figure 1.: SGIP and WatterSaver BLU Comparison

Figure 1 compares weekday load shapes for SGIP HPWH and WatterSaver BLU 4–9, showing how each approach shifts energy use relative to its baseline profile. WatterSaver BLU load begins to rise near the end of the peak period, while SGIP HPWH maintains a more consistent reduction through the shed window. One possible explanation is that more WatterSaver water heaters begin running short of hot water late in the peak period, causing the heat pump—or, in some cases, resistance elements—to turn on to maintain customer hot water service. This would reduce the observed load reduction during the shed hours. The SGIP HPWH and WatterSaver baseline shapes also differ slightly with WatterSaver showing lower baseline usage, likely reflecting differences in the metered populations and water heater control strategies.

Figure 2 presents the SGIP HPWH baseline and JA13 shifted load shapes for the sample of JA13 shifting water heaters and the load shape for WatterSaver ALU 4-9 baseline and shifted water heaters.

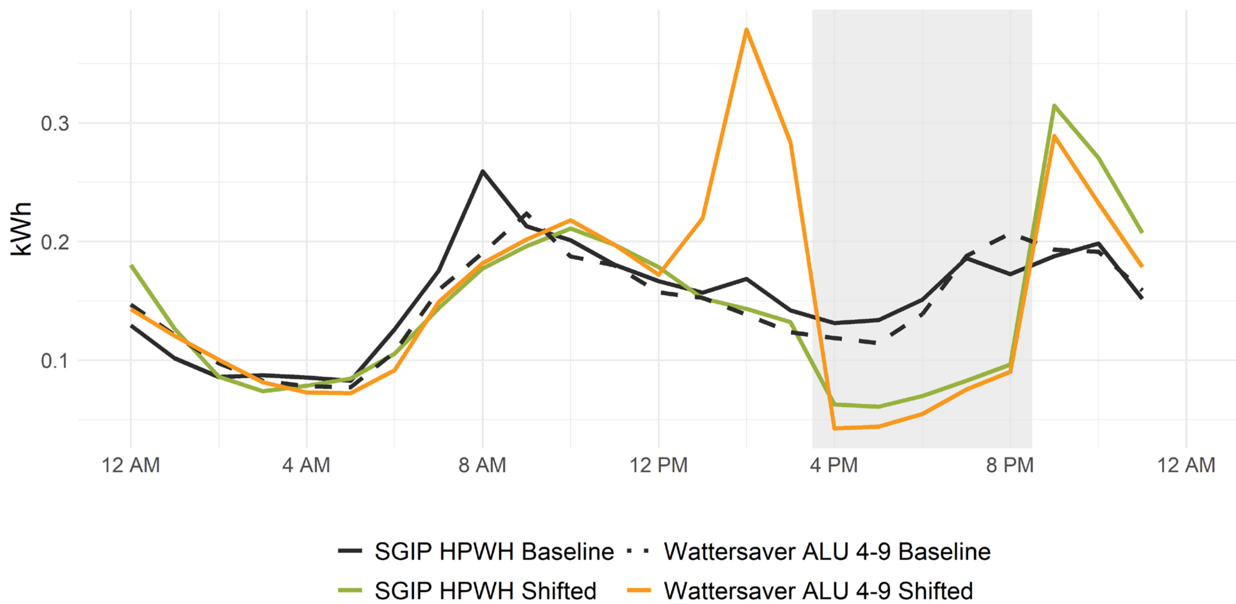

Figure 2.: SGIP HPWH and WatterSaver ALU Comparison

The ALU 4–9 load shape shows a pronounced increase before the shed window. This reflects ALU preconditioning: if possible, the water heater raises the tank temperature above the customer set point in the hours before the shed period, storing additional thermal energy that can be used during peak hours. BLU does not superheat the tank during load up, and SGIP HPWH JA13/TOU control does not include this preconditioning step. ALU therefore produces a larger shed impact, although part of the difference may also reflect baseline differences. As a result, comparisons across these groups should be considered preliminary and interpreted with caution. This finding also underscores the importance of using comparable baseline load shapes when comparing load-shifting strategies.

Conclusion

Signaled ALU currently produces larger shed-period reductions than scheduled TOU control, but with higher daily energy use. WatterSaver's ALU strategy reduces energy use more during shed hours and produces a smaller post-shed snapback than scheduled control. However, it also increases daily energy use and can slightly increase customer bills, depending on the applicable TOU price differential. The ALU algorithm has since been refined to improve daily energy and bill outcomes by making setpoint changes incremental rather than absolute and by limiting resistance-element operation during load-up and snapback. Scheduled control remains much less expensive—potentially nearly free—to implement, and it could evolve to include ALU features. One manufacturer is already incorporating ALU into JA13 logic for its newest generation of HPWHs.

Requiring scheduled TOU load shifting for new HPWHs could provide a low-cost pathway to broad load-shifting benefits. If scheduled load shifting were enabled for all new HPWHs, California could reduce peak demand without materially affecting customer comfort. The main incremental costs would likely be additional installer time and a more robust verification process. Scheduled control may also reduce customer bills: Verdant's interim

WatterSaver analysis found that BLU saved customers about $25 per year on average. Although these savings are modest for an individual customer, the combined peak-demand and bill-reduction benefits could be substantial across millions of HPWHs. Implementation would still need to address customer choice, including clear opt-out options.

Device-level non-compliance materially reduces realized impacts for both scheduled and signaled approaches. In the SGIP HPWH evaluation sample, 44% of devices were not following JA13 at the time of evaluation. This issue was traced to a contractor app flaw that appeared to confirm successful enrollment even when the schedule was not enabled. The manufacturer has since deployed a fix, although a post-update review found that some devices continued operating on custom user schedules that override JA13. In WatterSaver, 15–32% of commanded days resulted in failed shifts, depending on platoon; the longer 3 PM–12 AM shed window had the highest failure rate, partly because of participant opt-outs. Together, these results show that the gap between enrolled devices and effectively shifting devices is large enough to reduce program impacts. Standardized field verification of JA13 activation and transparent reporting of command failure rates would improve the reliability and measurability of both approaches.

Because HPWHs currently use relatively little energy during the 4–9 PM peak period, the available load-shifting potential is limited. Average HPWH consumption during current evening peak hours is relatively low, leaving less load available to shift. As California electrifies more buildings, however, winter morning peaks may become more important. If TOU periods begin to include morning peaks, they may align more closely with periods of higher HPWH usage. In that case, HPWHs could provide greater load-shifting value than they do under today's afternoon and evening peak periods.

References

Besson A., N. Dirr, and H. Liu, 2024, "Heat Pump Water Daily Load Shifting: Advanced Load Up and Evaluation Challenges" In 2024 ACEEE Summer Study on Energy Efficiency in Buildings, August 4-9, 2024, Monterey, CA, USA, Panel 1)

Besson A., S. Barsun, and H. Liu, 2026. "Hotter is Not Always Better – Impacts of HPWH Settings on Load Shifting Performance" In *2026 ACEEE Summer Study on Energy Efficiency in Buildings*, August 2-7, 2026. Monterey, CA, USA, Panel 1

Kehmeier, E., D. Urigwe, S. DeBauche, T. Weintraut, A. Besson, and S. Barsun. 2026. "Lessons from California's Large-Scale Deployment of Efficient, Flexible Heat Pump Water Heaters." Proceedings of the 2026 ACEEE Summer Study on Energy Efficiency in Buildings. Monterey, CA, Panel 2.

California Energy Commission (CEC), JA13 Qualification Requirements available at https://www.energy.ca.gov/sites/default/files/2025-11/JA13_2025_Qualification_Requirements_HPWH_DM_ADA.pdf

Hughes J., S. Barsun, S. Rosenbrock, A. Besson, H. Liu, 2025, "Turning Up the Heat (Before Peak): Lessons from a Heat Pump Water Heating Load-Shifting Pilot", In 2025 International Energy Program Evaluation Conference, Denver, CO

Verdant Associates, Self-Generation Incentive Program Heat Pump Water Heater Interim Evaluation, expected publication in summer 2026