



X1932: Demand Response EM&V Support Study

June 3, 2022

SUBMITTED TO:
Connecticut Energy Efficiency Board

SUBMITTED BY:
NMR Group, Inc., DSA, DNV

PREPARED BY:
Savannah Horner, DSA. Aaron Schrader, DNV.

Abstract

The X1932 Demand Response EM&V Support Study explored Connecticut’s demand response (DR) programs. The study reviewed three programs from the United Illuminating demand response portfolio, including two residential and one commercial program, as well as one residential Eversource program not examined as part of a multi-state DR evaluation.¹ In addition to producing impact estimates for these programs, the goal of the study was to assess the appropriateness of methods and ensure alignment with program design and evaluation goals. The study had three fundamental **findings**:

- Evaluations based on advanced metering infrastructure (AMI) data require fewer assumptions and produce more robust load impact estimates than analysis of telemetry from connected devices.
- Baseline methods that are well-documented improve replicability and transparency.
- Connecticut would benefit from standard processes to regularly estimate DR program impacts, report results, and incorporate those impacts into planning.

The study also includes **recommendations** based on the four evaluated programs, as well as overarching recommendations for how these DR programs may provide value to the grid and the implications of incorporating DR programs into the Program Savings Document (PSD):

Recommendation	Smart Savers Rewards	Wi-Fi HPWH	C&I Auto DR	Wi-Fi AC	Value of DR	PSD Implications
Revisit the connected load assumption.	✓					
Ensure all enrolled devices are dispatched.	✓	✓		✓		
Use AMI where available	✓	✓	✓	✓		✓
Target less efficient equipment that has coincident loads.		✓				
Use same-day event notification, after the adjustment period.			✓			
Establish a load predictability requirement.			✓			
Incentivize participants to inform the program of planned load decreases.			✓			
Modify the adjustment window if pre-event changes are deployed.	✓	✓	✓	✓		✓
Implement a clear settlement baseline methodology that is consistently applied.	✓	✓	✓	✓		✓
Track the association between device and utility account/meter number.	✓	✓				
Make curtailment algorithms more aggressive.				✓		
Assess data quality prior to including device manufacturers in the program.	✓	✓		✓		
Assess the possibility of utilizing other cost-effectiveness tests.					✓	
Consider bidding DR resources into the ISO-NE marketplace.	✓	✓	✓	✓	✓	
Define standard reporting methods/criteria for quantifying the value of DR.	✓	✓	✓	✓	✓	✓

¹ Several of Eversource’s Connecticut DR programs have been evaluated as part of joint efforts with Massachusetts.

Table of Contents

ABSTRACT	2
EXECUTIVE SUMMARY	1
KEY FINDINGS & RECOMMENDATIONS – DEMAND RESPONSE PROGRAMS	1
United Illuminating Smart Savers Rewards	2
United Illuminating Wi-Fi Enabled Heat Pump Water Heaters	3
United Illuminating C&I Auto Demand Response	4
Eversource ConnectedSolutions Wi-Fi Air Conditioners	5
KEY FINDINGS & RECOMMENDATIONS – VALUE OF DR AS A GRID RESOURCE	6
IMPLICATIONS FOR THE PSD	7
SECTION 1 ACTIVE DEMAND REDUCTION STRATEGIES IN 2019 – 2021 CONSERVATION AND LOAD MANAGEMENT PLAN	9
1.1 UNITED ILLUMINATING	9
1.1.1 Residential	9
1.1.2 Commercial	10
1.2 EVERSOURCE	10
1.2.1 Residential	10
1.2.2 Commercial	11
SECTION 2 UNITED ILLUMINATING SMART SAVERS REWARDS	13
2.1 PROGRAM OVERVIEW	13
2.2 EVENT SUMMARY	14
2.3 DATA	16
2.3.1 Enrollment	16
2.3.2 Runtime	18
2.3.3 AMI	18
2.4 METHODS	19
2.4.1 Baseline	19
2.4.2 Connected load	21
2.5 CONNECTED LOAD	21
2.6 EX POST IMPACTS	25
2.6.1 Residential	26
2.6.2 Small Business	31

2.7 RECOMMENDATIONS.....34

SECTION 3 UNITED ILLUMINATING WI-FI ENABLED HEAT PUMP WATER HEATERS35

3.1 PROGRAM OVERVIEW.....35

3.2 EVENT SUMMARY35

3.3 DATA36

 3.3.1 Enrollment Data36

 3.3.2 Runtime.....37

3.4 EX POST IMPACTS.....37

3.5 RECOMMENDATIONS.....38

SECTION 4 UNITED ILLUMINATING C&I AUTO DEMAND RESPONSE PROGRAM40

4.1 PROGRAM OVERVIEW.....40

4.2 METHODS42

 4.2.1 Data42

 4.2.2 Baselines42

4.3 IMPACT VERIFICATION44

4.4 IMPACT EVALUATION47

 4.4.1 Asset 1.....48

 4.4.2 Asset 2.....50

 4.4.3 Asset 3.....51

4.5 FINDINGS AND RECOMMENDATIONS52

SECTION 5 EVERSOURCE CONNECTEDSOLUTIONS WI-FI AIR CONDITIONERS55

5.1 PROGRAM OVERVIEW.....55

5.2 EVENT SUMMARY55

5.3 DATA56

 5.3.1 Enrollment.....56

 5.3.2 State Change.....57

5.4 METHODS58

 5.4.1 Converting to Electric Demand58

 5.4.2 Baseline59

5.5 EX POST IMPACTS.....59

5.6 RECOMMENDATIONS.....61

SECTION 6 VALUE OF DR AS A GRID RESOURCE62

6.1 INTRODUCTION.....62

6.2 OVERVIEW OF DEMAND RESPONSE.....63

6.3 ELECTRIC DEMAND RESPONSE IN CONNECTICUT.....65

6.4 BENEFITS AND COST OF DEMAND RESPONSE.....66

6.5 PROGRAMS IN OTHER STATES70

6.6 AVOIDED ENERGY SUPPLY COMPONENTS (AESC).....72

6.7 PROGRAM COST-EFFECTIVENESS METHODOLOGY73

6.8 ENERGY AND CLIMATE GOALS74

6.9 CONCLUSION.....78

6.10 RECOMMENDATIONS FOR FUTURE PROGRAM DESIGN79

SECTION 7 STUDY IMPLICATIONS FOR THE CONNECTICUT PSD.....80

7.1 COMPONENTS FOR CONSIDERATION IN PSD GUIDANCE80

7.1.1 Performance Assessment Process.....81

7.2 REVIEW OF OTHER JURISDICTIONS82

7.3 RECOMMENDATIONS.....84

APPENDIX A UI SMART SAVERS BASELINE ACCURACY.....85

APPENDIX B UI C&I AUTO DR REGRESSION SPECIFICATIONS.....91

Executive Summary

This report presents the results from the X1932 Demand Response (DR) EM&V Support Study. The study reviewed three programs from the United Illuminating demand response portfolio (Smart Savers Rewards, Wi-Fi Enabled Heat Pump Water Heaters, and the C&I Auto Demand Response Program), as well as one from the Eversource demand response portfolio (ConnectedSolutions Window A/C Controls). In total, these programs represented approximately 3 MW of demand response capability during summer 2021 but are expected to grow over time. The objectives of the study were to produce defensible estimates of kW savings for these programs, assess the appropriateness of program designs, and make recommendations to improve the evaluability of DR programs in Connecticut. The study includes several recommendations on how to increase per-participant impacts and identifies methodological issues that should be resolved so the utilities can make informed decisions about whether increased participation is a wise investment.

KEY FINDINGS & RECOMMENDATIONS – DEMAND RESPONSE PROGRAMS

- **Evaluations based on advanced metering infrastructure (AMI) data require fewer assumptions.** Each of the residential DR programs evaluated faced data availability and data quality challenges with the telemetry supplied by device manufacturers. Runtime data from connected devices also requires assumptions about equipment capacity and efficiency that are not necessary when analyzing hourly or sub-hourly data from the utility revenue meter.
- **Baseline methods that are well-documented improve replicability and transparency.** The evaluation team found that the electric distribution companies (EDCs) and their implementation contractors tended to use a 10-of-10² baseline with symmetric additive adjustment. This approach is the default settlement baseline for the ISO-NE energy market, so it is a reasonable status quo method in Connecticut. The evaluation team recommends that Connecticut specify the adjusted 10-of-10 baseline as the default DR baseline, but includes a discussion of applications where alternate methods may be needed such as:
 - **Sites with behind-the-meter solar PV.** Solar generation is an unknown, variable driver of the net load shape, which is difficult to control for with the settlement baseline.
 - **Programs that notify participants of DR events more than one hour prior to the beginning of the event.** A notification prior to the one-hour adjustment window allows participants to game during this window, which can influence the additive adjustment.

² A 10-of-10 baseline is the simple average of hourly (or sub-hourly) loads for the ten most recent non-event weekdays.

- **Programs that employ pre-cooling or pre-heating strategies that begin more than one hour prior to the beginning of an event.** A pre-cooling or pre-heating strategy prior to the one-hour adjustment window artificially influences the additive adjustment.
- **Connecticut would benefit from a clearly defined demand response framework.** As DR programs mature and expand in Connecticut, it will be important to clearly define the policy objectives, performance metrics, valuation assumptions, reporting expectations, and evaluation paradigm. The Conservation and Load Management Plan and Program Savings Document are the two logical venues to memorialize this type of information.

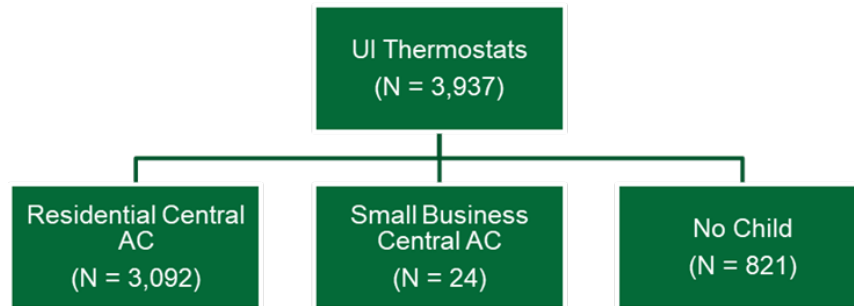
United Illuminating Smart Savers Rewards

The Smart Savers Rewards program is a connected thermostat program. During event hours, UI creates events in the EnergyHub Demand Response Management System (DRMS) and the DRMS communicates with thermostat manufacturers to issue control the program thermostats' temperature set-points by shifting the set-point up to four degrees from the initial temperature. Using a 10-of-10 baseline, this evaluation produced ex post impacts for summer 2019, 2020, and 2021. Additionally, it assessed the accuracy of the connected load assumption. These objectives are accomplished using thermostat-level runtime data (telemetry) and whole-building interval data (AMI). A set of program recommendations were developed based on the findings in this evaluation.

- **Recommendation 1: Use AMI data where it is available.** The quality of runtime data varies widely by device manufacturer, which creates caveats when using this data for evaluation purposes. Using data from United Illuminating's AMI network instead would allow for a more straight-forward calculation of program impacts. Using runtime data is industry standard for utilities that don't have AMI, but most utilities with AMI leverage their AMI investment to measure DR impacts.
- **Recommendation 2: Revise the connected load assumption to 2.1 kW per thermostat.** The current assumption used to convert AC runtime, 3.5 kW, is too high for Connecticut and leads to overstated cooling load and ultimately the DR impacts for this program. Given the limitations that are presented when discussing the connected load calculations, this connected load assumption should be further studied in the future.
- **Recommendation 3: Require thermostats to be assigned a child group for dispatch.** When thermostats are registered to the Smart Savers program, they are placed into groups based on their characteristics. At first, all thermostats are placed into a parent group of "UI Thermostats". Based on the customer's classification, thermostats should then either be placed into the "Residential Central AC" or the "Small Business Central AC" child group. As seen in [Figure 1](#), there is a subset of thermostats that are not placed into either of the defined child groups, which means these thermostats are not dispatched when events are called at the child group level. Allowing thermostats to sit without a child group creates an equity issue, since enrolled thermostats receive an incentive regardless of how often they were dispatched during a DR season. It also limits DR performance, as there is a subset of enrolled thermostats that are not being dispatched during DR events.

During summer 2021, approximately 20% of thermostats were not dispatched for seven of the nine events. Since the program is expected to grow significantly 2022-2024, preserving the ability to dispatch residential and commercial separately makes sense.

Figure 1: Thermostat Hierarchy



United Illuminating Wi-Fi Enabled Heat Pump Water Heaters

The Wi-Fi-enabled HPWH program dispatches demand response events in both the summer and winter, by triggering the HPWH manufacturers’ demand response curtailment algorithm. The goal of this evaluation was like that of the Smart Savers Rewards program. However, the data available for this evaluation included only the estimated consumption on event days and the exact curtailment algorithm (how HPWH operation was manipulated) was unclear. Without consumption data on non-event days, constructing the baseline, and ultimately the load reduction, is not possible. Nevertheless, the demand impacts reported in the post-event summary showed small reductions of 0.03 kW per-device for the years with non-zero reporting, while the Smart Savers program reduced load by 0.77 kW per device during this same period. This is a product of the following aspects:

- 1 Because HPWHs are highly efficient there is very little electric load available to curtail for DR.
- 2 Customers may have already enabled the "Eco Mode" that the DR event dispatches.
- 3 Residential water heating load is not coincident with the afternoon/evening peaks.

With these small load reductions, the study recommends that the program:

- **Recommendation 1: Target less efficient equipment.** The high efficiency of HPWHs severely limits the DR potential of the program. Even with the expansion of more devices, this program will be limited in aggregate impacts.
- **Recommendation 2: Target equipment with coincident loads.** Since residential water heating load is highest in the morning and summer and winter peaks occur in the afternoon and evening, respectively, direct load control of water heaters provides less value in the in Connecticut than a system with winter peaks in the morning hours.
- **Recommendation 3: Do not enroll additional devices from manufacturers who cannot supply quality telemetry.** Reliable hourly or sub-hourly device-level operating data should be a pre-requisite for inclusion in any connected device demand response program. Telemetry data allows program administrators to determine which devices are responding, troubleshoot dispatch issues, and report preliminary load impacts.
- **Recommendation 4: Leverage the storage capability of water heaters.** Tank water heaters can be used as a thermal storage device to store clean energy from solar during the day and reduce ramping in the afternoon and evening.³ If pre-heating is used, it is important that there is a cold water mixing mechanism to prevent scalding.

United Illuminating does not plan to target HPWH for DR after 2022 once it's contract with the current residential DR vendor ends. UI will continue to seek out other DR opportunities in the water heating field with less efficient equipment.

United Illuminating C&I Auto Demand Response

Working closely with UI and OEM A, C&I customers select pre-programmed demand response strategies that go into effect during event hours. This evaluation assessed the methodology and load reduction associated with the three sites that were enrolled in the program during 2019. Ultimately, it will always be difficult to accurately estimate load reduction for a small number of sites with highly variable load. However, of the baselines tested, the standardized mean 10-of-10 with adjustment provides the best load reduction estimate. Note the evaluated load reduction is substantially different than reported values. Since the program only had three participating sites in 2019 and no events were called in 2020, this evaluation looks closely at individual site results from 2019 to arrive at the following recommendations:

- **Recommendation 1: Implement a clear settlement baseline methodology that is consistently applied.** This will increase fairness and transparency, particularly for the settlement baseline methodology for sites with solar generation, which was not well-defined nor systematically replicable.
- **Recommendation 2: Use a same-day event notification that occurs after the adjustment period.** Same-day notification can improve baseline adjustments by decreasing the probability of event-related load effects or gaming during the adjustment window. If participants are notified of an event prior to the interval(s) used to adjust the baseline, there is a chance of influencing loads (up or down) by knowledge of and

³ https://www.dret-ca.com/wp-content/uploads/2021/04/pge_wattsaver_field_test_dret_final_report.pdf

preparation for the upcoming event. Since the demand response strategies are automated, we believe the potential benefit of a more accurate baseline greatly outweighs the minimal risk of additional program burden and/or dissatisfaction on participating customers.

- **Recommendation 3:** The program should modify the adjustment period if load reduction strategies involve pre-event changes to load (e.g., pre-cooling). For example, Massachusetts C&I demand response programs use the hour beginning 2 hours prior to the event as the adjustment window. This adjustment window decreases the likelihood that pre-event changes to load are included in the baseline adjustment. Further, the longer adjustment period can also help to decrease the impact of a spurious change in load leading to a poor baseline adjustment. ISO-NE provides same-day event notification and uses the 15-minute interval preceding event notification as the adjustment window. If the program changes to a same-day event notification, this is also another possibility.
- **Recommendation 4: Establish a load predictability requirement.** Successful assessment of load reduction for a DR program, either for settlement or ex post evaluation, requires a balance of site-level load predictability with the expected magnitude of load reduction.
- **Recommendation 5: Define how to handle the presence of solar, since it is a substantial issue for the measurement of DR.** For sites with solar generation, the program should obtain the solar production data to reconstitute site load and then estimate customer baseline load and load reduction.
- **Recommendation 6:** Build on the pilot's level of customer participation and event frequency to create a grid resource that is more reliable and more evaluable than the limited 2019 pilot activity. Greater numbers of participants can be expected to provide more reliable load reduction and load reduction estimates, whereas smaller population sizes, such as that seen in this study, can be prone to the nuances of one or two customers having an outsized influence on the overall result. Addressing the additional challenges presented by the demand response automation process, as noted earlier, may naturally increase enrollment rates and the number of participants.

Eversource ConnectedSolutions Wi-Fi Air Conditioners

Eversource began the ConnectedSolutions Window A/C Controls program in 2020 for customers that are otherwise unable to participate in the Wi-Fi thermostat program due to a lack of central A/C. The program implementer, ThinkEco, enrolls and orchestrates DR events for Wi-Fi enabled equipment to reduce cooling loads. This evaluation assessed the accuracy of the employed 10-of-10 baseline and estimated ex post impacts for summer 2021 events. Based on the findings in this evaluation, we developed a set of program recommendations.

- **Recommendation 1: Assess device manufacturer's data quality prior to their inclusion in the program.** The quality of telemetry data varies by manufacturer. The absence of accurate telemetry creates uncertainty regarding the program performance. If a device manufacturer is unable to provide this data, there is no assurance that demand

response events are being dispatched to the customer and no way to quantify their load reductions.

- **Recommendation 2: Ensure device manufacturers provide guidance on when a unit is drawing power.** Even with adequate state-change data, assumptions must still be made about the power draw of the unit. Without clarification from the manufacturers, the assumption implemented could either be over-stating or under-stating the load of these devices.
- **Recommendation 3: Make curtailment algorithms more aggressive.** Window AC and mini-splits have small peak loads and on average, these devices reduce electric load for the cooling end use by 38%, as compared to the 58% reduction from the Eversource Wi-Fi thermostat program⁴. The Wi-Fi thermostat program controls central ducted systems rather than Window AC units, which may allow for more aggressive curtailment. The ConnectedSolutions program can't affect equipment size and reference load. If program economics are strained with average kW impacts of 0.12 kW per device, Eversource might explore with the equipment manufacturers curtailing cooling operations more aggressively to produce larger load reductions. More aggressive control can result in attrition due to participant discomfort during events so finding the right curtailment strategy is a balancing act for the utility.

KEY FINDINGS & RECOMMENDATIONS – VALUE OF DR AS A GRID RESOURCE

The value of demand response lies in the ability to cost-effectively manage supply and demand on the grid with direct control or market signals. Historically, the value of demand response has been in its ability to provide load relief when the grid is constrained. The utilities are maximizing the avoided cost benefits allowed under the current state designated cost-benefit framework. As markets evolve and technologies advance, demand response programs and the cost-effectiveness frameworks they operate within need to evolve as well. For example, FERC 2222 allows increased participation for DERs, making it increasingly important for resources such as DR to perform during dynamic peak periods that may be driven by new system loads such as electric vehicles and heat pumps. The additional value streams created by the flexibility of distributed resources, (e.g., wholesale market payments), are in the benefit cost analyses often found in other jurisdictions, but not being used in Connecticut.

- **Recommendation 1: We recommend the utilities encourage the possibility of utilizing a different cost-effective test, which is a state level decision.** Monetizing and incorporating additional value streams including avoided environmental and compliance costs as seen in other states can boost program cost-effectiveness and help meet state policy goals like decarbonization and enhanced environmental quality. Regulators direct utilities in other states to monetize or use proxy values to account for a variety of environmental externalities. Considering additional avoided costs that reflect reduced environmental impact or increase human wellbeing—in the primary test—can bring forth

⁴ <https://ma-eeac.org/wp-content/uploads/2019-Residential-Wi-Fi-Thermostat-DLC-Evaluation-Report-2020-04-01-with-Infographic.pdf> Table 4-6

identification of potentially new value streams but do require a benefit cost framework that allows them.

- **Recommendation 2: Consider bidding DR resources into the ISO-NE market.** By not participating in the ISO-NE market, UI and Eversource forego a base payment for availability and a pay-for-performance payment when ISO-NE calls on resources to reduce load. Instead, program economics rely on the premise that peak shaving will lower their peak load forecast and future capacity obligations. If the EDCs opt to pursue wholesale recognition of their DR programs, there are ways to mitigate participation risks (e.g., using qualified Curtailment Service Providers or aggregators) to make the risk profile of participating more acceptable. Note that FERC 2222 requires ISO and distribution companies to coordinate DER participation in both markets⁵.

IMPLICATIONS FOR THE PSD

While the value of demand response evaluations as ad hoc processes to determine program impacts is clear, there is more value in defining standard processes to regularly estimate program impacts and incorporate those impacts into utility planning. In Connecticut, the Program Savings Document (PSD) provides guidance on how to calculate savings for energy efficiency programs but does not discuss processes for assessing demand response savings. The study recommends that Connecticut:

- **Recommendation 1: Standardize the reporting methods and criteria for quantifying the capability and value of demand response programs.** Some trends emerge from our review of other jurisdictions. First, only programs that are bid into the RTO/ISO market are subject to estimation methodologies defined by ISO rules. Second, most states that have PUC-defined load impact protocols or evaluation frameworks produce annual impact evaluations. Third, methods or estimates defined by TRMs tend to be simpler than other approaches, where deemed savings are used or a scalar adjustment to deemed savings is used. Although the EDCs' demand response programs are not recognized at the wholesale level, the evaluation team still sees value in documenting some basic definitions, methods, and assumptions for demand response programs in the PSD.

ANONYMIZATION OF MANUFACTURER-SPECIFIC RESULTS

The UI Smart Savers Rewards and Wi-Fi Enabled Heat Pump Water Heater programs rely on a single vendor (EnergyHub) to communicate with and control devices from a variety of equipment manufacturers. Similarly, the Eversource ConnectedSolutions Wi-Fi Air Conditioners program implementer (ThinkEco during summer 2021) orchestrates curtailment of Window AC and ductless mini-split units manufactured by three different companies. The EDCs' contracts with these original equipment manufacturers (OEMs) include various provisions regarding public disclosure of market shares or proprietary demand response capabilities. This report employs an

⁵ ISO New England submitted the compliance filing on in February.

anonymization strategy for all references to equipment manufacturer. We use the term OEM A, OEM B, OEM C etc. in all discussions of manufacturer-specific findings. The lettering scheme is consistent throughout (e.g., “OEM A” always indicates the same connected thermostat manufacturer) and the EDCs have the anonymization key so that OEM-specific findings can be addressed, as appropriate.

Section 1 Active Demand Reduction Strategies in 2019 – 2021 Conservation and Load Management Plan

The Connecticut Conservation and Load Management (C&LM) Plan⁶ is part of a multi-year effort to deliver cost-effective energy efficiency and demand management programs to customers. Priority six of this plan, and the focus of this study, deals with the implementation of effective demand reduction strategies, which curtail load during seasonal demand peaks. Active demand reduction strategies are dispatchable, or event-based, in contrast to the passive “every day” peak demand reductions generated by energy-efficiency programs. Active demand reduction strategies are relatively new offerings for United Illuminating and Eversource. Pilots conducted from 2017 – 2019 gradually transitioned into full-fledged program offerings in 2019 and 2020.

Both United Illuminating and Eversource show projected costs and peak demand reductions, but the 2019-2021 C&LM Plan does not estimate the monetary benefits from the implementation of these active demand reduction strategies. Peak demand reductions can reduce capacity obligations, lower capacity costs, reduce stress on the electric grid, and mitigate energy price spikes. The Companies’ active demand reduction offerings are not currently recognized at the wholesale level as resources by ISO New England, so the underlying valuation premise is that lower peak load will reduce the capacity obligations and associated costs going forward. [Section 6](#) contains a detailed discussion of the economics of active demand reduction strategies in Connecticut.

1.1 UNITED ILLUMINATING

United Illuminating (UI) offers various active demand reduction, or demand response, programs to its customers, which include three residential and one commercial offering. Across these programs, we evaluate the following programs: Smart Savers Rewards, the Wi-Fi Enabled Heat Pump Water Heaters, and the C&I Auto Demand Response. Although we are not evaluating the Peak Time Rebate Pilot, a description of the program is included to align with its inclusion in the Conservation and Load Management Plan.

1.1.1 Residential

- **Smart Savers Rewards (Bring Your Own Thermostat).** The Smart Savers Rewards program is a connected thermostat program, which began in the summer of 2019 and continued through the summer of 2021. It is a direct load control program that is comprised of residential and small business customers who enroll their own thermostats. During event hours, UI controls these thermostats’ temperature set-points by shifting the set-point up to four degrees from the initial temperature. Each event lasts two hours, and upon completion, the thermostat returns to its typical operations. As an incentive, customers

⁶ <https://portal.ct.gov/-/media/DEEP/energy/ConserLoadMgmt/FINAL-2021-Plan-Update-Filed-10302020.pdf>

receive \$25 per thermostat they enroll and another \$25 per thermostat at the end of each summer season.

- **Wi-Fi Enabled Heat Pump Water Heaters.** The Wi-Fi-enabled HPWH program began in 2018 and continued through 2021. This program dispatches demand response events in both the summer and winter, which triggers the HPWH manufacturers' demand response curtailment algorithm. Thus far, demand reductions have been limited due to the highly efficient nature of the HPWHs, the low number of enrolled devices, and the lack of weather-dependency, which makes targeting peak usage more difficult.
- **Peak Time Rebate Pilot [Cancelled].** This pilot began in the winter of 2018 and continued through the summer of 2019. It included both a home energy report (HER) and an incentive, in the form of a rebate, to reduce energy use during event hours. UI's HER component generated savings similar that of other HER programs, while the active demand response strategy produced minimal savings on event days.

1.1.2 Commercial

- **Auto Demand Response Program.** In 2020, UI transitioned their Auto Demand Response Pilot into a fully operational demand response program. Qualifying participants were defined as C&I customers who had either an Energy Management System, a Building Management System, a lighting control system, or a Programmable Logic Controller, as well as had a curtailable load they were willing to eliminate or reduce during event hours. Working closely with UI and OEM A, customers selected pre-programmed demand response strategies that would go into effect during event hours. As an incentive, customers enrolled in the program can receive up to \$50 per reduced kW, with a minimum commitment of 50kW of load reductions for a three-year period.

1.2 EVERSOURCE

Much like UI, Eversource offers multiple demand response programs to its customers. The study only evaluates the ConnectedSolutions Window A/C Controls program. The other Eversource offerings are evaluated as part of a multi-state effort.

1.2.1 Residential

- **ConnectedSolutions Window A/C Controls.** Launched in 2020, Eversource began the ConnectedSolutions Window A/C Controls program for the segment of homes unable to participate in the Wi-Fi thermostat program due to a lack of central A/C. Eversource controls these devices directly by making slight adjustments to the temperature setting or turning the air conditioner on and off periodically to lower energy use during times of peak demand. With the success of the Connected Wi-Fi Thermostats program, the main challenges of this program surround incentivizing commitment and continued participation.
- **Connected Wi-Fi Thermostats and HVAC Systems (Bring Your Own Device).** In 2019, Eversource transitioned their Bring Your Own Device pilot into a fully operational demand response program, continuing through the summer of 2021. It is a direct load control

program in which customers enroll their own devices, which are controlled by Eversource during event hours. The most recent evaluation found that each controlled thermostat reduced peak demand by 0.81 kW⁷.

- **Battery Storage.** Based on findings from the 2018 and 2019 Massachusetts battery storage projects, Connecticut deployed a battery storage pilot in 2020. During this roll-out year, there were 170 batteries enrolled⁸. Given this level of enrollment, Eversource decided to continue the program and expand to include new battery partners in 2021.
- **Electric Vehicle (EV) Charger Control.** The growing load and the flexibility of electric vehicle charging in Connecticut present an ideal end use for demand response. With 70 residential charging stations enrolled, Connecticut began their electric vehicle charger control program in 2020.
- **Direct Communication to the EV.** The goal of this program is to develop an electric vehicle load shifting system, which will lead to reduced costs for electric vehicle customers. For this initiative, the system will collect data surrounding driving and charging behaviors while customers receive notifications designed to encourage charging at more optimal times.

1.2.2 Commercial

- **C&I Active Demand Response Strategies.** The Eversource active demand response strategies for commercial customers remain in the pilot phase during the 2021 season. The pilot has currently produced findings regarding the need for different technologies across differing dispatch strategies.

Table 1 shows the 2021 projections of active demand reduction capability by Company and sector presented in the 2021 update to the 2019 – 2021 C&LM Plan. The Eversource DR portfolio comes mostly from the C&I sectors while approximately 75% of UI’s DR capability is residential.

Table 1: Projected 2021 DR Capability (kW) by Company and Sector

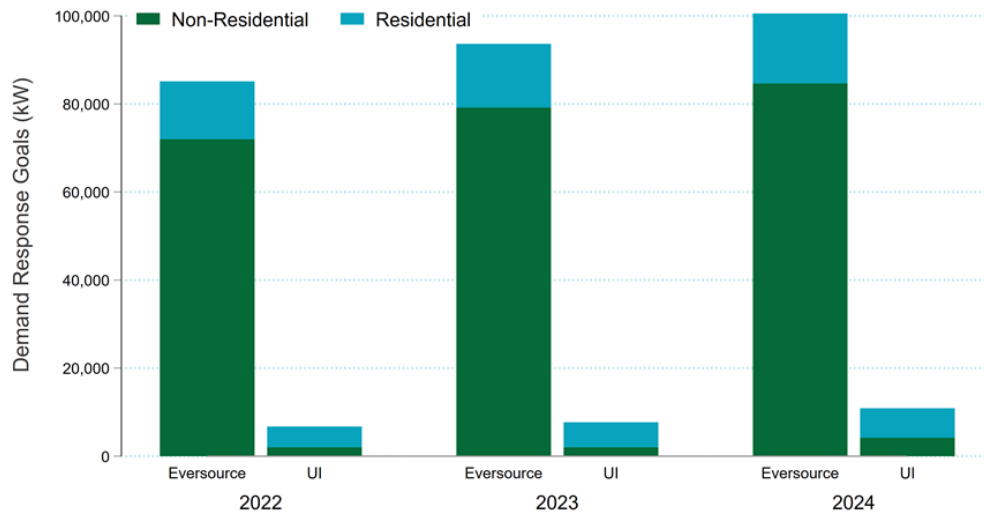
Company	Residential Demand Response	C&I Demand Response
Eversource	6,839	36,807
United Illuminating	4,480	1,498

In March 2022, the utilities filed a proposed C&LM Plan for 2022 – 2024 with the demand response goals shown in Figure 2.

⁷ <https://ma-eeac.org/wp-content/uploads/2019-Residential-Wi-Fi-Thermostat-DLC-Evaluation-Report-2020-04-01-with-Infographic.pdf> Table 4-6

⁸ <https://portal.ct.gov/-/media/DEEP/energy/ConserLoadMgmt/FINAL-2021-Plan-Update-Filed-10302020.pdf>

Figure 2: Proposed 2022-2024 C&LM Demand Response Goals



Section 2 United Illuminating Smart Savers Rewards

Key Findings

- Telemetry data quality varies widely by device manufacturer.
- During summer 2021, approximately 20% of thermostats failed to be dispatched for 7 of the 9 events.
- The 3.5 kW connected load assumption used to convert AC runtime is too high for Connecticut and should be revised to 2.1 kW.

2.1 PROGRAM OVERVIEW

This section summarizes the evaluation of United Illuminating (UI) Company's Smart Savers Rewards connected thermostat demand response program. This program began in the summer of 2019 and has continued through the summer of 2021. It is a direct load control program that is comprised of residential and small business customers who enroll their own thermostats.

EnergyHub, the implementation contractor selected by UI, controls the thermostats' temperature set-points by shifting the set-point up to four degrees from the initial temperature during the event window. Prior to the event, thermostats are also dispatched for pre-cooling of the house. Each event lasts two hours, and upon completion, the thermostat returns to its typical programming. In addition to the direct control, UI also communicates with customers to better understand and control their energy usage.

This evaluation investigates the accuracy of the baseline and the connected load, as well as discusses the ex post impacts. These objectives are accomplished using two data sources:

1. **Telemetry.** Thermostat-level runtime data, in 15-minute intervals, delivered from the EnergyHub portal.
2. **AMI.** Whole-building hourly interval data, in kWh, from UI's advanced metering infrastructure.

Using these data sources, the study compares the runtime reductions from the telemetry to the load reductions from the AMI data to construct an estimate of the connected load, or the runtime to power conversion, for smart thermostats in UI's service territory. This connected load is then used to quantify demand reductions by using non-event day consumption patterns to estimate what customer loads would have been on the event days had they not been dispatched.

[Section 2.3](#) will discuss our data sources and [Section 2.4](#) will discuss our methodology for estimating baseline accuracy, connected load, and ex post impacts. Finally, [Section 2.6](#) will discuss results, while [Section 2.7](#) will offer recommendations.

2.2 EVENT SUMMARY

Across the 2019, 2020, and 2021 seasons, there were 19 events called. The event start times varied between 3 PM and 5 PM, lasting two hours. Table 2 details the maximum temperature during the event⁹, the number of thermostats targeted, and the event start and end time for each event. Additionally, those days that ISO-NE identified as the peak day for each month are denoted with an asterisk. Across the three years, UI dispatched demand response (DR) events on six of the nine ISO-NE peak days. Of these six peak days, all DR events covered the peak hour except for the August 19, 2019, event. This event spanned from 4 PM to 6 PM, but the peak hour occurred at 3 PM. The occurrence of this peak hour coincides with the pre-cooling dispatch of the program, which means this program contributed to the peak load during the peak hour. This outcome is worse than not dispatching a DR event.

Table 2: Event Day Characteristics

Event Date	Max Event Temperature (F)	Thermostats Targeted		Event Start	Event End
		Residential	Small Business		
7/17/2019	81	2,270		4 PM	6 PM
7/19/2019	85	2,270		5 PM	7 PM
7/30/2019*	87	2,306		4 PM	6 PM
8/8/2019	84	2,323	21	4 PM	6 PM
8/19/2019*	82	2,357	22	4 PM	6 PM
8/29/2019	81	2,372	23	4 PM	6 PM
6/23/2020*	81	812		5 PM	7 PM
7/9/2020	84	773	11	5 PM	7 PM
7/20/2020	91	2,157	21	5 PM	7 PM
8/11/2020*	83	2,043	23	5 PM	7 PM
6/7/2021	79	3,969		5 PM	7 PM
6/8/2021	79	3,890		3 PM	5 PM
6/21/2021	77	2,810	18	5 PM	7 PM
6/29/2021*	87	2,804	18	5 PM	7 PM
8/11/2021	83	3,092	18	5 PM	7 PM
8/12/2021*	88	3,092	13	5 PM	7 PM
8/13/2021	90	2,670	18	4 PM	6 PM
8/25/2021	86	3,080	24	5 PM	7 PM
8/26/2021	87	3,080	24	5 PM	7 PM

* ISO-NE System Load Peak Day

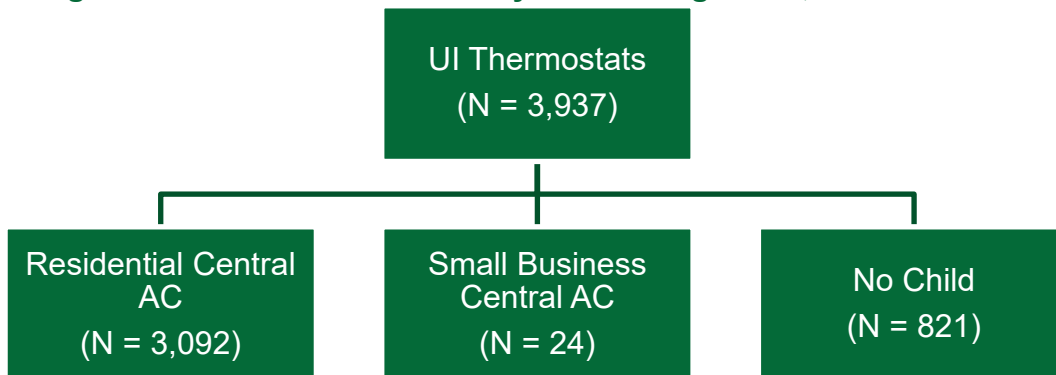
⁹ Uses the six closest Connecticut weather stations from the enrolled thermostat file to construct the average temperature across the region.

The small business and residential thermostats are always dispatched at the same time on event days. When the two groups are dispatched jointly, the EnergyHub reporting delivers the performance of all thermostats, without a distinction between the small business and residential thermostats. In most cases, they are independently dispatched as two separate events, except during the beginning of each season where they are dispatched jointly.

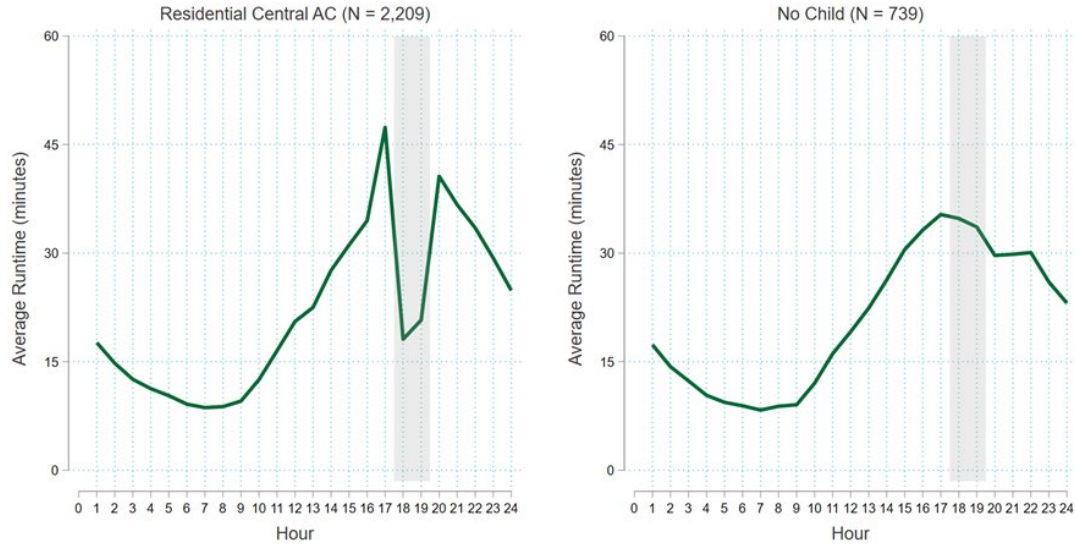
Across the three demand response seasons, the number of enrolled thermostats grew, but there is a large dip in dispatched thermostats during the first two events of the 2020 season. During these two events, none of the enrolled OEM C thermostats were eligible for dispatch, due to configuration changes. This large drop in available thermostats will be important to consider when evaluating the impacts and connected load assumption using the 2020 data.

In addition to these dispatch disparities across thermostat manufacturers, there are also dispatch disparities between thermostat groups. When thermostats are registered to the Smart Savers program, they are placed into groups based on their characteristics. At first, all thermostats are placed into a parent group of “UI Thermostats.” Based on the customer’s classification, thermostats should then either be placed into the “Residential Central A/C” or the “Small Business Central A/C” child group. Figure 3 displays this hierarchy for the August 12, 2021, event. In addition to the “Residential Central A/C” and the “Small Business Central /AC” groups, there are 821 thermostats without a child group. Those thermostats that are exclusively contained in the “No Child” group are only dispatched on the June 23, July 9, and July 20 events in 2020 and the June 7 and June 8 events in 2021. With the limited amount of dispatch dates, there is untapped DR potential.

Figure 3: Thermostat Hierachy for the August 12, 2021 Event



The left side of Figure 4 shows the average performance of the thermostats in the “Residential Central AC” group on the August 12, 2021, event date. This group was dispatched from 5 PM to 7 PM and clearly displays the runtime reductions associated with a DR event. The right side shows those thermostats in the “No Child” group on the same day. It is clear that these thermostats were not dispatched.

Figure 4: Residential Central AC and No Child on the August 12, 2021 Event

The percent of thermostats in the “No Child” group, and inevitably the percent of thermostats excluded from the dispatch strategy, is growing with the recent commencement of the UI Uplight marketplace program. Thermostats are brought into this program through the UI Marketplace¹⁰ and initially placed in the “No Child” group. With a subset of thermostats not being immediately placed into a child group, DR performance is limited. It is imperative that these thermostats sitting in the parent group be moved into their respective child group to maximize the DR performance of this program.

2.3 DATA

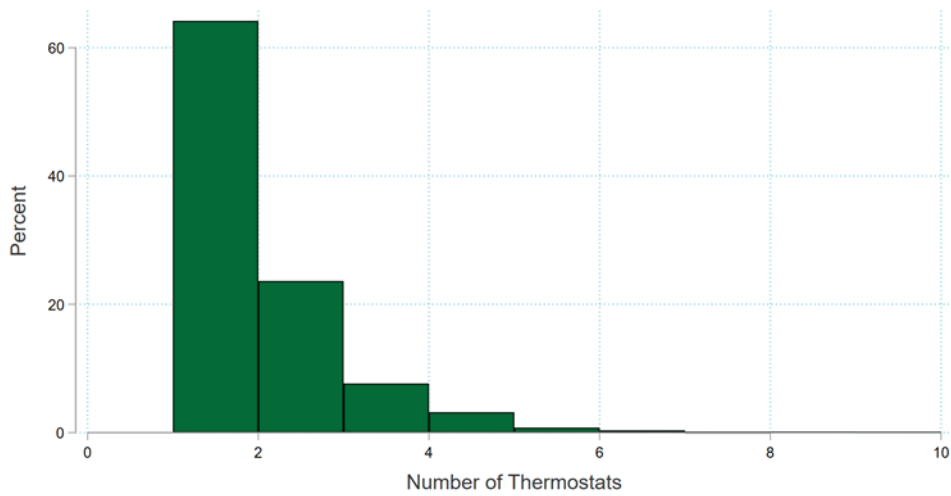
2.3.1 Enrollment

The enrollment data contains key pieces of information about each participant, including their account number, location, and enrollment start and end dates. Since the runtime and AMI data came from two separate files, the information in this enrollment file, specifically a unique identifier, is necessary to link the two datasets. Unfortunately, of the 4,887 unique thermostats in the enrollment file, 26.2% have missing account numbers. Due to the incomplete account mapping, we used the combination of service address and service first and last name as a proxy for a given account.

Using this proxy, the evaluation team assessed how many thermostats were associated with each account. Of all unique thermostats that were enrolled at some point, the average number of thermostats per proxy account is 1.6. This distribution is displayed in [Figure 5](#).

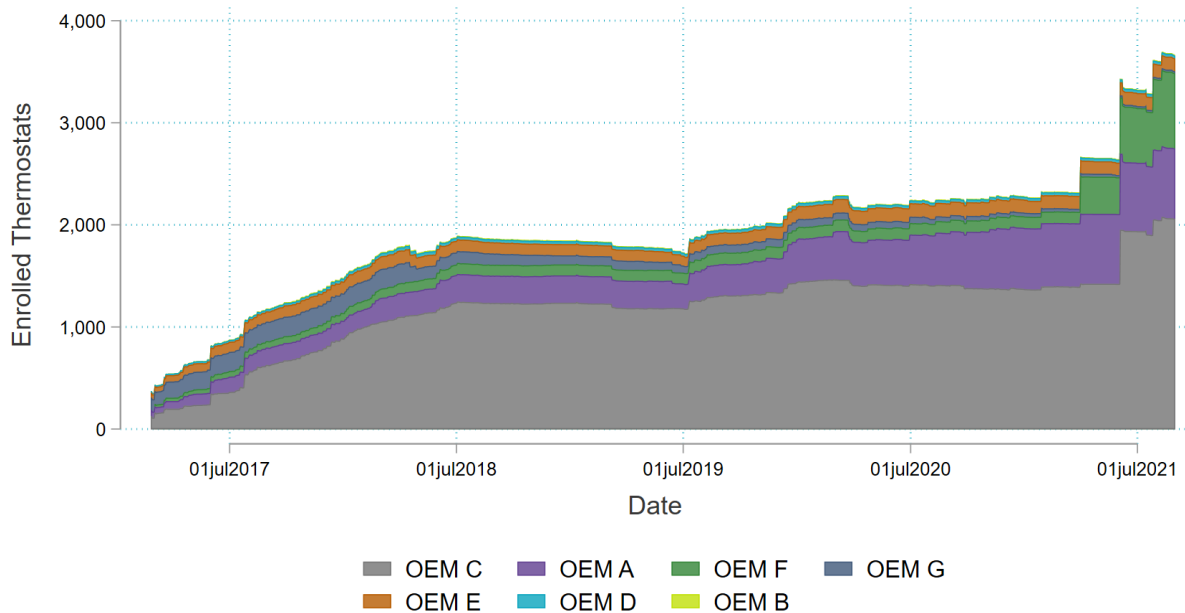
¹⁰ <https://uismartsolutions.com/>

Figure 5: Distribution of Thermostats at a Service Address



It is also important to understand that these 4,887 thermostats were not all enrolled at the same time. Figure 6 shows the number of actively enrolled thermostats by manufacturer from 2017 to the end of the 2021 DR season. While the number of enrolled thermostats has grown over time, OEM C consistently represents most thermostats. The next largest set of enrolled thermostats are manufactured by OEM A.

Figure 6: Enrolled Thermostats by Day



From this, we drew a sample of 270 unique residential participants to request AMI data. In lieu of using the unique account number to draw the sample, we again used the unique combination of service first and last name, service address, and service postal code, to identify accounts. These

sample accounts have a range of thermostats associated with a given premise (from 1 to 8, with a mean value of 1.8), include all thermostat manufacturers, and cover most of the zip codes indicated in the enrollment file. Of the 270 unique residential participants, the account number was missing for 45 (17%). We provided an additional 45 participants that had account numbers, if UI was not able to identify the AMI meters associated with the requested participants through our unique combination of characteristics.

2.3.2 Runtime

The thermostat runtime files provide the number of minutes during each 15-minute interval that the HVAC system is in each operation mode for each thermostat. It also provides the operation mode of the unit during each interval, indoor temperature, outdoor temperature, and thermostat set-point values. These files were pulled from the EnergyHub portal for the 2019, 2020, and 2021 DR seasons.

There are two types of thermostat runtime files: full-summer and event-day. The full-summer runtime files span June through August, while the event-day runtime files only record data on event days. Due to the limited number of dates in the event-day runtime files, they should only be used for supplemental information, such as the number of minutes the thermostat participated in the DR event for each interval.

The runtime data we downloaded was incomplete and came with many caveats. For the 2019 season:

- The full-summer runtime file recorded zero runtime for all thermostats.
- The event-day runtime only recorded non-zero runtime during event hours for OEM E and OEM G thermostats.

These gaps in the 2019 data are due to runtime data for participating thermostats only being stored for one year. Since this data was pulled in 2021, the full 2019 data was no longer available. Additionally, the 2020 and 2021 seasons had data quality issues, including:

- OEM E and OEM G thermostats recorded zero runtime across the entire DR season.
- OEM A thermostats were excluded from the full-summer runtime file and only reported runtime from 2 PM to 9 PM in the event-day runtime file. With only event day data present, a baseline cannot be constructed for this thermostat manufacturer.

These data collection deficiencies¹¹ reduced the number of thermostats that are eligible to be aggregated into our runtime analysis.

2.3.3 AMI

The AMI file contains the whole-house hourly electric consumption data at the account level for the 2019 and 2020 DR seasons. We requested AMI data for 270 accounts and received AMI data for 231 of them. Using the enrollment file, we attempted to match these accounts with those that

¹¹ These issues are present in other jurisdictions but were particularly prevalent for UI during the period of analysis. It's not exclusively an EnergyHub, or OEM, or UI issue. There are a lot of complex data streams to manage and requires a lot of hands-on vigilance.

we requested. [Table 3](#) details the differential we experienced between merging these two files. We can confidently gather account details, such as the number of thermostats and the location, for the 158 accounts that match on account number. The other 73 accounts only have a partial match to the enrollment file, and we cannot ascertain participant attributes.

Table 3: AMI Linkages to the Enrollment File

Match Type	Number of Accounts
Account Number	158
Name and Address – No account initially provided	37
Name and Address – Wrong account initially provided	24
Address	9
Not in documentation	3
Total	231

Additionally, for the 231 accounts that we received AMI data, only 29 can be linked to usable runtime. This is most notably due to the thermostat manufacturer runtime reporting summarized in [Table 4](#).

Table 4: AMI Linkages to the Runtime Files

Thermostat Manufacturer	Correspondence Issues	# of Accounts
OEM C	Anonymized in telemetry	139
OEM A	Only provides telemetry between 2 PM – 9 PM on event days	44
OEM F	-	23
OEM E	Records all zero runtime	10
OEM G	Records all zero runtime	6
OEM D	-	4
OEM B	-	2
Total Accounts in Both		29

Without sufficient linkages between the runtime and the AMI file, we are unable to analyze the same subset of customers across the two file types. Although this limitation is present, the 231 customers for whom we received AMI data still constitutes a representative sample of the participants, with an average 1.8 thermostats per account.

2.4 METHODS

2.4.1 Baseline

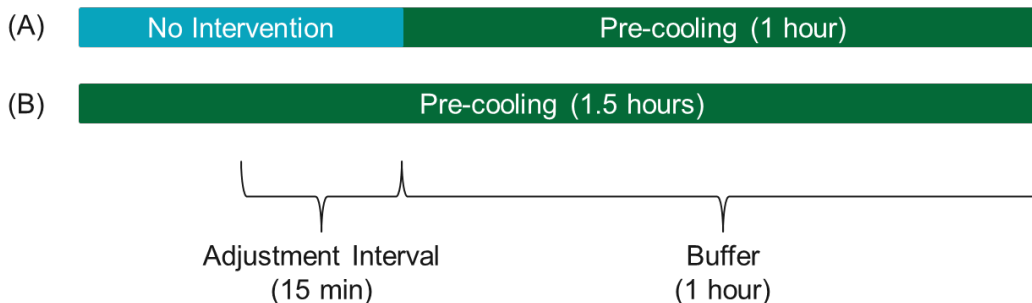
To estimate demand reductions, it is necessary to estimate what electric load would have been on an event day in the absence of a DR event. This estimated load, commonly referred to as a counterfactual, can be approximated using a baseline. We rely on non-event day consumption

patterns to estimate our baseline. Using these non-event days allows us to model how customers would have consumed energy on an event day had they not been dispatched. Baselines are simply a tool to produce demand reductions estimates, not the end itself. They help filter out noise and explain variation, to allow the demand reduction to be more easily detected.

The baseline that EnergyHub uses to construct impacts is the default ISO-NE settlement baseline. This is a mean 10-of-10 baseline¹² with an additive adjustment¹³, which is based upon an hour buffer and a 15-minute pre-event adjustment interval¹⁴. The last ten eligible days are defined as non-holiday, non-event weekdays. The study included a baseline accuracy assessment on this ISO-NE baseline (defined in the Baseline Accuracy appendix). This exercise leads the evaluation team to believe that the ISO-NE baseline performs reasonably well and should continue to be implemented given the simplicity and regional acceptance.

The one-hour buffer implemented in this baseline happens to coincide with the 60 minutes of pre-cooling prior to the commencement of the event, Figure 7(A). Constructing an adjustment that is based upon the pre-cooling period would lead to an inaccurate, inflated baseline, which could become an issue if pre-cooling is extended beyond one hour in this program, Figure 7(B). Ultimately, this would lead to higher impacts during the event period because the baseline is adjusted up too much. We recommend UI clarify with EnergyHub and the OEMs that pre-cooling must be limited to one hour as the ISO-NE 10-of-10 baseline is used to estimate performance.

Figure 7: Pre-Cooling and Adjustment Interval Relationship



The change in energy use, or the impact, is calculated as the difference between the baseline and the observed consumption during the event. We calculate this impact on a per-thermostat basis for the runtime data and a per-home basis for the AMI data, which is subsequently scaled down to a per thermostat basis. To reconstruct the EnergyHub average load shed value, these

¹² Simple average of the ten most recent non-event, weekdays.

¹³ The symmetric additive adjustment scales the simple 10-of-10 baseline up or down based on the ratio of the event day loads to the baseline loads in the hours prior to the event. In practice, this adjustment typically scales the baseline up to account for the fact that event days are hotter than the non-event days used to construct the baseline.

¹⁴ Many baselines of this type have a day-of adjustment that scales the baseline to match the observed day-of loads in an adjustment interval that occurs at some point before the event begins. The intention of a pre-event buffer is to accommodate customer notification time or pre-cooling while also preventing that period from being used as part of the adjustment window. Including periods where a customer is aware of the upcoming event in the adjustment window risks customer or vendor manipulation of loads to increase the perceived reduction during the event (by adjusting the baseline upwards).

per-thermostat impacts are averaged across the two-hour event window and then scaled by the number of targeted thermostats.

2.4.2 Connected Load

One of the critical assumptions when developing impact estimates from runtime data is the connected load assumption. This value is used to convert cooling runtime to cooling load (kW). This runtime to power conversion represents the average power draw of the cooling equipment when operating. The telemetry files only show cooling runtime, in minutes, which implies that the AC unit is either on or off. In the case of newer variable speed compressors, this assumption may be violated because the compressor might run for 60 minutes in an hour, but at 50% capacity.

Since EnergyHub reporting uses thermostat runtime data to estimate the load impacts in kW, we want to assess the conversion factor from time to energy on a per-thermostat basis. When using the AMI data, we divide the total estimated kW reduction by the number of thermostats to get a per-thermostat estimate. Generally, we think of the relationship between thermostat and AC unit as one-to-one, but there are also zoned systems where two thermostats control a single AC unit.

Once the AMI data is scaled down, the runtime must be converted from minutes to the percentage of time the unit is running. In the equation below, we assume that the runtime is in hourly intervals; to convert to the percentage of time, we divide by 60 minutes.

$$\text{Cooling kW} = \text{Connected Load} * \frac{\text{Minutes of cooling runtime}}{60}$$

Ultimately, to solve for the connected load, we use the calculated runtime reductions and AMI impacts across the event window, the hour prior to the event (pre-cooling), and the hour after the event (snapback). Across this set of hours, we select the connected load assumption that minimizes the sum of squared errors. The sum of square error quantifies how much of the variation in observed values cannot be explained by the model in question. The lower the SSE, the better the fit. For example, this equation would allow us to calculate a connected load of 3 kW, if there is an estimated runtime reduction of 30 minutes, or 0.5 hours, and a demand reduction of 1.5 kW during a DR event.

Since both the runtime and demand reductions are required to solve for the connected load, we are only able to use the 2020 DR season. This is because runtime is unavailable in 2019 and AMI data was not requested for 2021. Additionally, in 2020, the AMI data is for a random sample of customers and the runtime is not a census, due to thermostat manufacturer nuances. These issues indicate that it is certainly possible for the average home in our AMI sample to be a different size than in our runtime data sample. With that being said, the average number of program thermostats per home in the program as a whole and in the AMI sample is 1.6 and 1.8, respectively.

2.5 CONNECTED LOAD

The connected load assumption used by EnergyHub for UI is a 3.5 kW draw for the average cooling system. [Table 5](#) shows the connected load used for an example event day in each DR season, in which Hour 1 indicates the reading at the beginning of the first hour of the event and

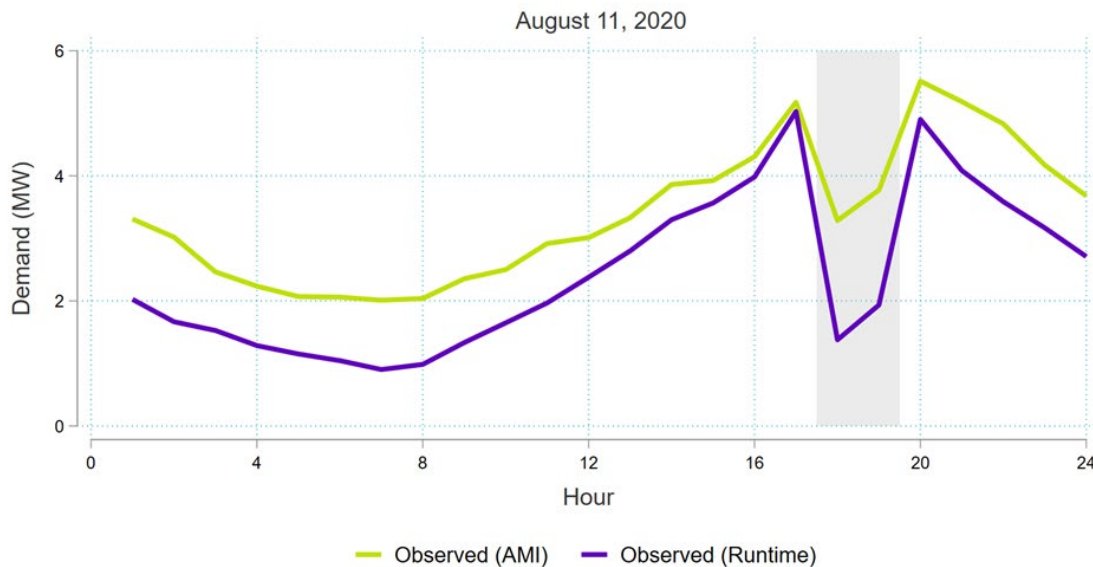
Hour 2 indicates the reading at the beginning of the second hour of the event. It is evident that between 2019 and 2020 the connected load assumption increased from 3.2 to 3.5 kW, while 2021 maintained the same assumption as 2020.

Table 5: Connected Load from the EnergyHub Portal

	July 17, 2019		June 23, 2020		June 7, 2021	
	Hour 1	Hour 2	Hour 1	Hour 2	Hour 1	Hour 2
kW Reduction	0.95	0.93	0.77	0.78	1.04	0.99
Runtime Reduction	0.30	0.29	0.22	0.22	0.29	0.28
Connected Load	3.17	3.21	3.50	3.56	3.57	3.55

When plotting the runtime and AMI on the same graph for the 2020 event days, it became clear that 3.5 kW per hour of runtime was overstating this assumption. In [Figure 8](#), we can see that at 5 PM the observed AMI and the observed runtime are almost identical. Given that the AMI is on a whole-house level and the runtime is on a device-level, these two loads would only be the same if no other device in the household was using any electricity for that entire hour. Ultimately, the device-level runtime should always be lower than the whole-house AMI, since the whole-house AMI includes additional electric load, including lighting and household appliances. Generally, this figure indicates that the 3.5 kW assumption overstates cooling load and, by extension, the DR reduction.

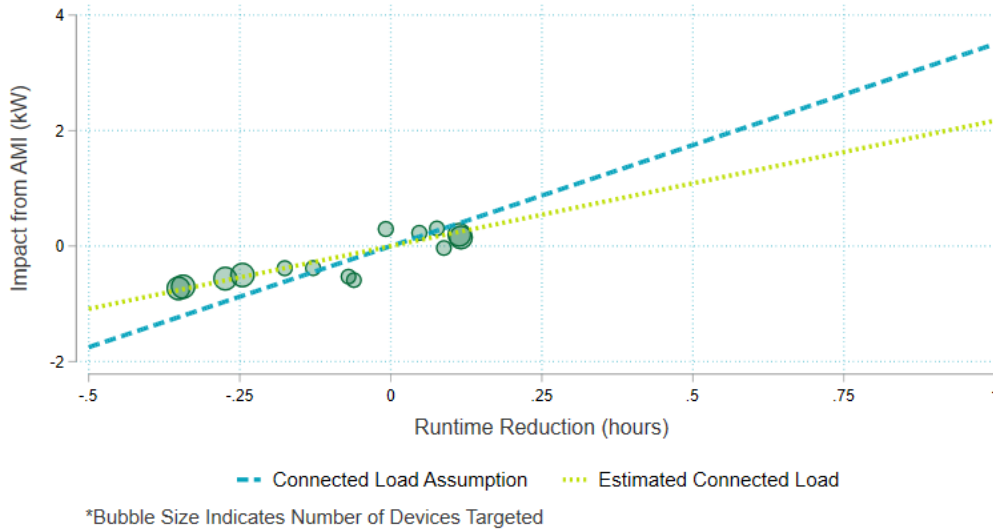
Figure 8: Connected Load Assumption on August 11, 2020 Event



Analyzing the runtime reductions and comparing them to the per-device kW impacts optimizes the connected load assumption. In [Figure 9](#), we plot each hour in the four-hour event window (one pre-cooling hour, two event hours, and one hour of post-event snapback) for each of the four

events in the 2020 DR season¹⁵. The blue line depicts the current connected load assumption of 3.5 kW. The slope of the line of best fit (green line) demonstrates the relationship between the runtime reductions and kW impacts, or the estimated connected load. It appears that a more accurate connected load is below the original assumption of 3.5 kW, but it likely also greater than 2.0 kW.

Figure 9: Connected Load Assumption



To identify a good estimate for the connected load, we want to find a connected load that minimizes the SSE. To do that we loop through all possible connected loads between one and four and calculate the sum of square error produced by each runtime to power conversion factor. Seeing as though a large portion of thermostats were unable to be dispatched on the June 23, 2020, and July 9, 2020, event days, we weight the sum of square error by the number of thermostats targeted. This places more emphasis on those two event days where all thermostats were eligible for dispatch. Using this methodology, a connected load of 2.05 kW is found to minimize the sum of square error between the runtime data and the AMI data. The 2.05 kW conversion factor is used throughout this report to convert our runtime estimates to kW demand.

This calculated connected load assumption is 41% lower than the 3.5 kW assumption that EnergyHub utilizes. Due to this large differential, we sought out additional information regarding the typical capacity of cooling units in Connecticut. The 2019 Connecticut Residential Appliance Saturation Survey (RASS) gathered detailed information on the capacity and efficiency of residential condensing units in Connecticut.¹⁶ During this study, onsite visits were conducted to collect the performance specifications for a sample of cooling units. We specifically look at the cooling capacity for central air conditioning and air source heat pumps.

The units that were analyzed within the RASS were spread across various household types, including single and multi-family. The Smart Savers Rewards enrollment file we received has no

¹⁵ The July 20 event only includes the two hours of the event window because of an issue with a substantial number of thermostats being offline surrounding the event hours.

¹⁶ The 2019 RASS database is available at <https://app.box.com/s/dq78xqs166xlyfax8qk4aw625fetiicx>

direct indication concerning the size of the household that an account is connected to. As a proxy for household size, we look at the number of thermostats connected to an account. With some accounts having up to ten thermostats, the enrollment file suggests the presence of multi-family households within the program. Given the likely presence of single and multi-family households in the program, we look at cooling units across all housing types in the RASS.

Table 6 details the specifications across the three types of cooling units. Using the kW draw calculated from the cooling capacity and SEER, and weighted by count of each cooling type, we create a weighted average of the kW to construct an estimate of the connected load. These calculations produce an estimate of the connected load equal to 2.14.

Table 6: Connecticut RASS Cooling Specifications

Cooling Type	Count	Cooling Capacity (BTU/hour)	SEER	kW
Central Air – Packaged	5	20,040	12.0	1.67
Central Air – Split	71	28,919	12.9	2.25
Air Source Heat Pump	24	25,000	13.0	1.92
Weighted Average	--	27,534	12.85	2.14

The 2019 RASS did not collect information on EER ratings, which measures the efficiency of units at 95 degrees (F), but we estimate the corresponding weighted average EER would be approximately 11.2. Using these proxy EER values to estimate connected load returns an estimate of 2.46 kW.

This exercise supports our findings that using a 3.5 kW connected load assumption to convert runtime to power overstates the impacts of the residential component of the program. The 2.14 kW per thermostat connected load assumption based on the Connecticut RASS is slightly largely than the 2.05 kW connected load assumption we calculated through the conversion of runtime to kW, but these two estimates are much closer to each other than they are to the current 3.5 kW conversion factor.¹⁷

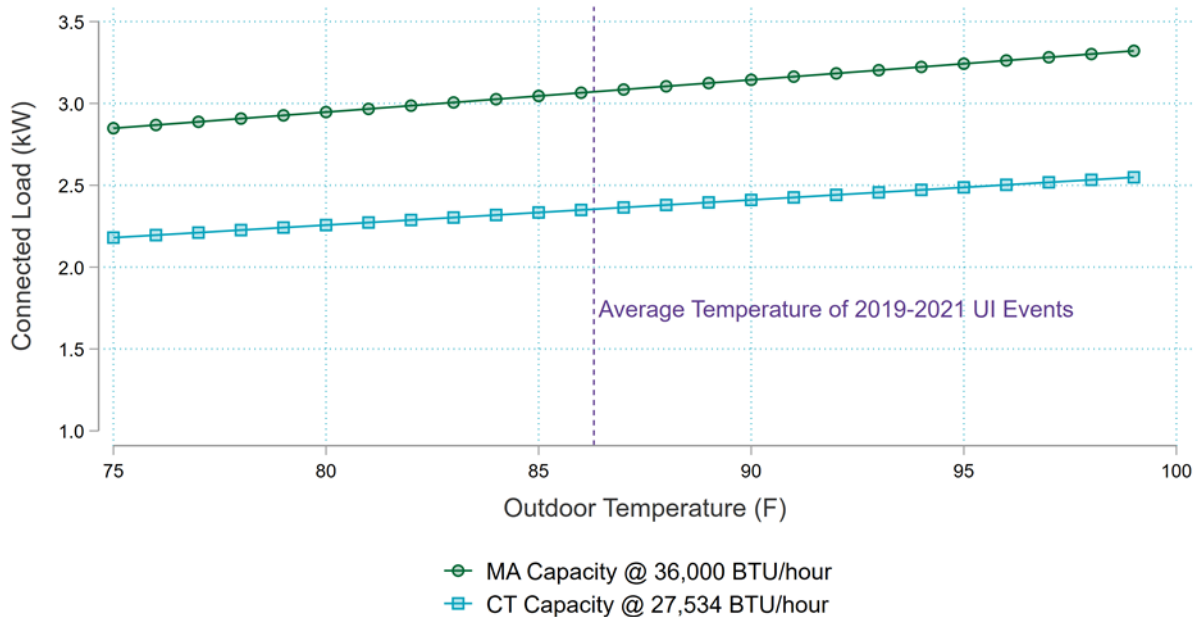
For another regional benchmark on this topic, the study team reviewed the results of 2019 Massachusetts Residential Wi-Fi Thermostat Direct Load Control Offering Evaluation.¹⁸ Absent AMI data, this study relied an assumed cooling capacity of 3 tons, or 36,000 BTU/hour,¹⁹ an EER of 10.7, and a linear equation relating kW draw to outdoor air temperature based on prior metering results. Figure 10 shows the Massachusetts connected load assumptions by outdoor air temperature as well as the weighted average cooling capacity from the 2019 RASS using the Massachusetts equations and efficiency assumptions. Even at 100 degrees (F), the Massachusetts approach does not reach 3.5 kW.

¹⁷ The 821 come from the marketplace and are associated with residential UI accounts. Commercial condensing units tend to be larger than residential condensing units. If the program grows the commercial side, a separate connected load assumption for commercial should be developed and used.

¹⁸ <https://ma-eeac.org/wp-content/uploads/2019-Residential-Wi-Fi-Thermostat-DLC-Evaluation-Report-2020-04-01-with-Infographic.pdf>

¹⁹ The basis of the 36,000 BTU/hour average cooling capacity assumption in Massachusetts is not reported

Figure 10: Comparison with Massachusetts Connected Load Assumptions

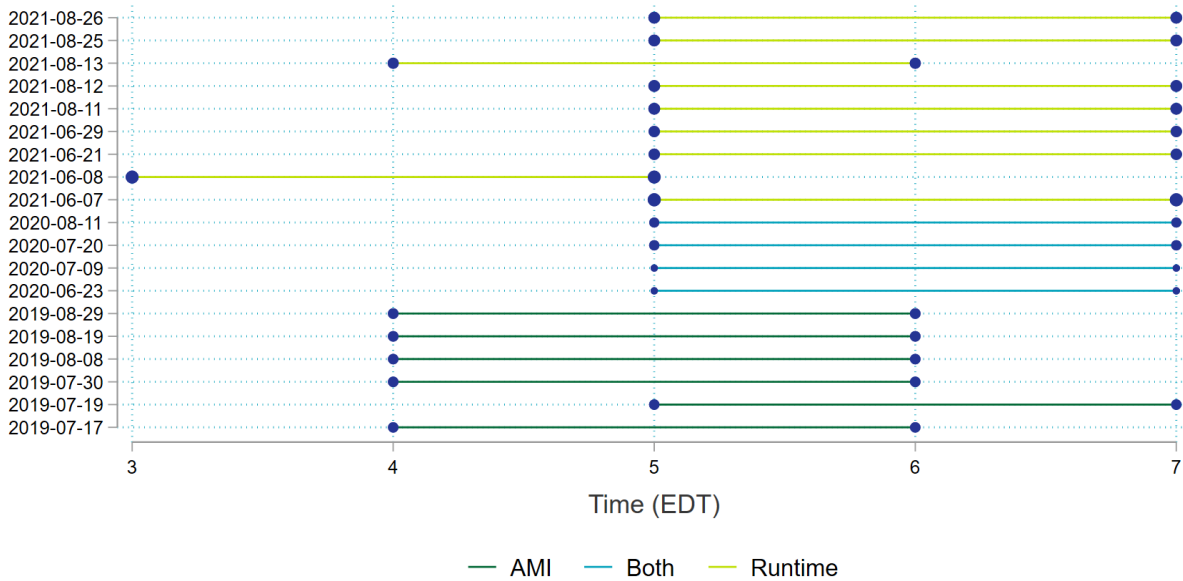


It is important to recognize the limitations this study faced when independently calculating the connected load assumption using the runtime and AMI data. The 2020 DR season only included four events, two of which were limited in the number of thermostats that could be dispatched. With these limitations, the most salient conclusion that can be drawn from this evaluation is that the assumption used to convert runtime to power should be much lower than the 3.5 kW conversion factor currently being used. Limitations aside, this conclusion is further substantiated by the 2019 RASS sample that indicates a 2.46 kW conversion factor and the assumptions used for similar offerings in Massachusetts. The study team recommends revising the connected load assumption to 2.1 kW for residential thermostats. The EER derived value comes from a RASS study, which is a random sample of CT households. The 2.1 kW value comes from the actual participants of this program. Given the limitations that are presented, this connected load assumption should be further studied moving forward. The study team did not investigate the connected load assumption for commercial thermostat due to the limited number of non-residential participants in the UI Smart Saver Rewards program. Typically, commercial AC units are larger than residential AC units. Absent further analysis of this issue, we recommend maintaining the 3.5 kW connected load assumption for the Small Business program component.

2.6 EX POST IMPACTS

The ex post analysis for 2019 and 2021 uses a different combination of data based on what was available for evaluation purposes. Figure 11 details what was used for each event, as well as the duration of the event window.

Figure 11: Event Duration and Datasets

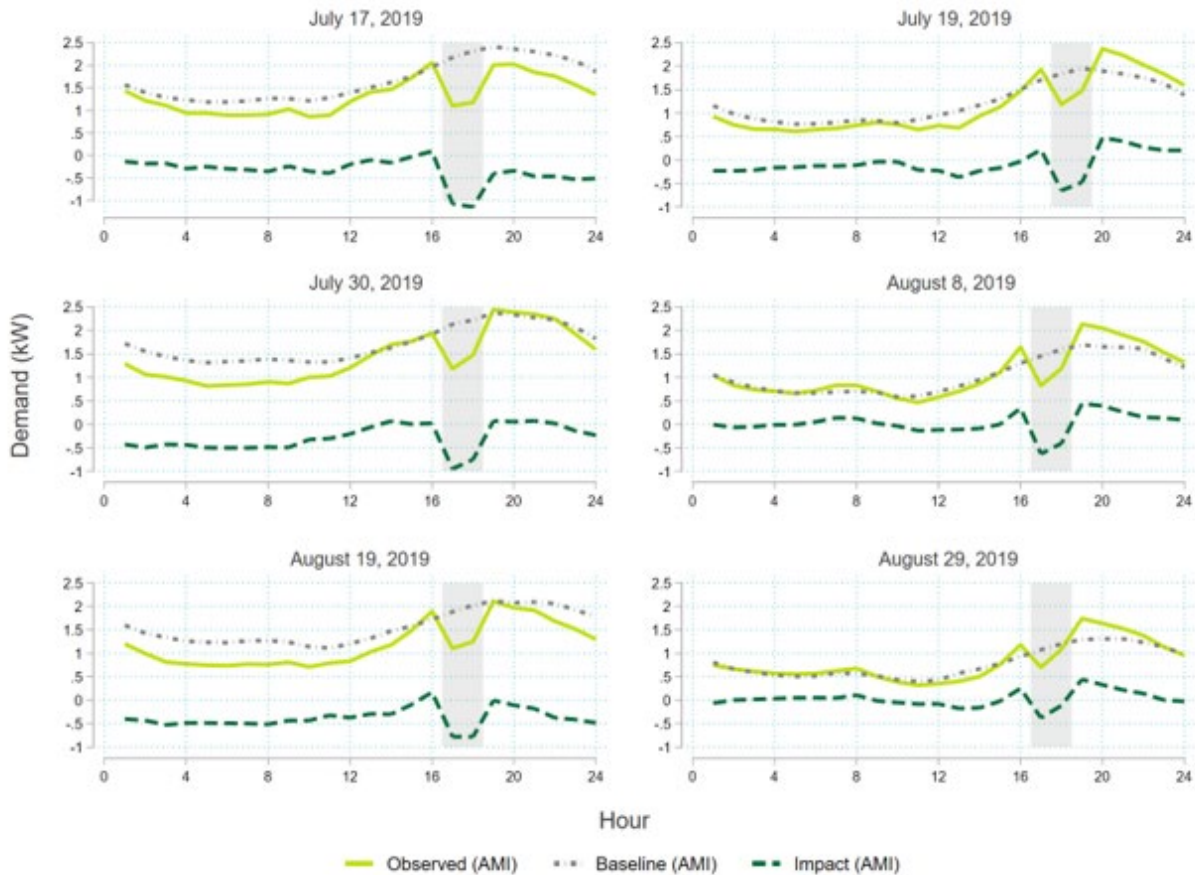


* Size of bubbles represents number of participants dispatched

2.6.1 Residential

For the six events in 2019, we are only able to analyze impacts using the sample accounts in the AMI data. As a representative sample, these accounts are used to produce the per-thermostat impacts in Figure 12. Each event day displays a clear drop in demand during the set of event hours.

Figure 12: Residential 2019 Event Days



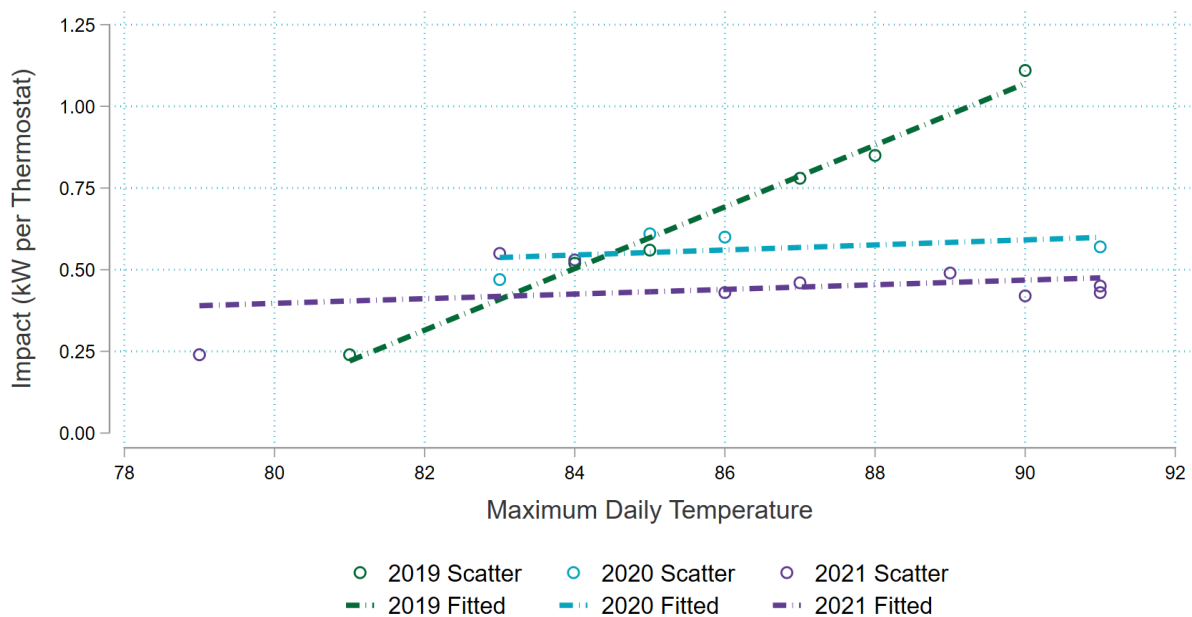
To understand these impacts at a program level, we use the average of the impacts produced during each event window and scale them by the number of thermostats targeted during the event. Across these six events, the average aggregate load shed calculated from the AMI data is similar to the estimates produced by EnergyHub, displayed in Table 7. In 2019, EnergyHub used a smaller connected load assumption (3.2 kW). This improved the convergence compared to 2020 when EnergyHub used a 3.5 kW connected assumption.

Table 7: Impacts of 2019 Residential Events

Date	Avg. Shed Per Thermostat (kW)		Aggregate Shed Across Targeted Thermostats (kW)	
	EnergyHub	AMI	EnergyHub	AMI
7/17/2019	0.91	1.11	2,070	2,509
7/19/2019	0.67	0.56	1,510	1,280
7/30/2019	0.80	0.85	1,840	1,956
8/8/2019	0.66	0.52	1,540	1,199
8/19/2019	0.84	0.78	1,990	1,834
8/29/2019	0.40	0.24	950	577
Average	0.71	0.68	1,650	1,559

Across the 2019 season, four of the six event days have an EnergyHub reported savings higher than the AMI evaluated savings. The July 17, 2019, and the July 30, 2019, events seem to be pulling the average load shed up, as seen in Figure 12. This is likely due to the combination of a set of mild baseline days and high temperatures in the hours prior to the event being called. The symmetric additive adjustment uses the ratio of event day loads to baseline loads prior to the event to scale the baseline up or down during the event hours so these factors work in tandem to create a large day-of adjustment. This influence is greatest on the July 17, 2019, event because the temperature drastically dropped during the event hours, due to an incoming storm. Additionally, Figure 13 helps to highlight that the 2019 impacts are more sensitive to temperature than the 2020 and 2021 events. This is likely a function of COVID-19 impacts, since more people stayed home during the summer of 2020 and 2021 and may have adjusted their thermostat during event hours, and thermostat dispatch inequities for both thermostats manufacturers and UI thermostat groupings.

Figure 13: Temperature and Savings Impact Relationship

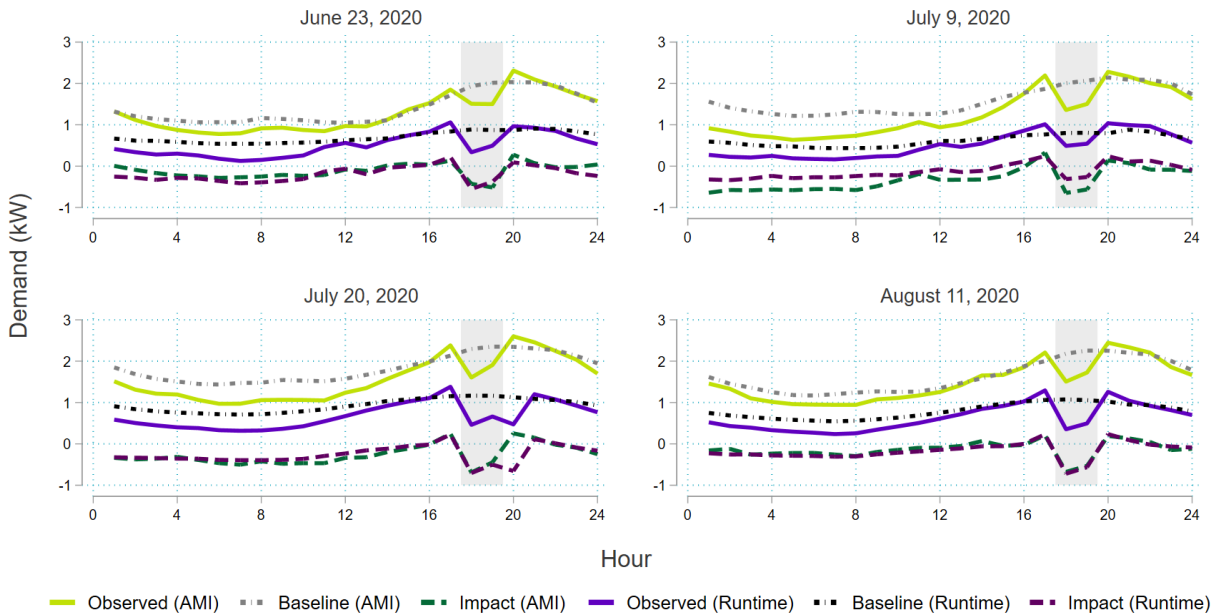


The mechanisms that create these inflated baselines should also be present in the EnergyHub reporting, but they are not. Without full insight into how the baseline was calculated in 2019 and/or without runtime data for comparison, we are unable to draw a conclusive opinion about why these two dates return such large impacts. However, the use of a 3.2 kW connected load assumption appears to return better estimates than the 3.5 kW assumption used in 2020 and 2021. It is unfortunate the study team was only able to get both runtime and AMI for one summer (2020) and that summer had just 4 events and two of them had vendor dispatch issues.

In 2020, we have both the runtime and AMI. Figure 14 graphs the per-thermostat impacts based on both datasets. This graph implements a connected load of 2.1 kW to convert the runtime to demand. With this connected load assumption, the impacts of the AMI and the impacts of the

runtime should be equivalent. The August 11, 2020, event day shows this relationship between the two impacts. Similarly, the July 20, 2020, event shows these equivalent impacts, except during the snapback hour, where a significant portion of thermostats were offline in the runtime data. The two events where OEM C thermostats were unavailable for dispatch display more deviation from this connected load assumption. The deviation on June 23, 2020, is not large and is consistent with a significant portion of thermostats not being eligible for dispatch. A more evident discrepancy is visible on the July 9, 2020, event day.

Figure 14: Residential 2020 Event Days



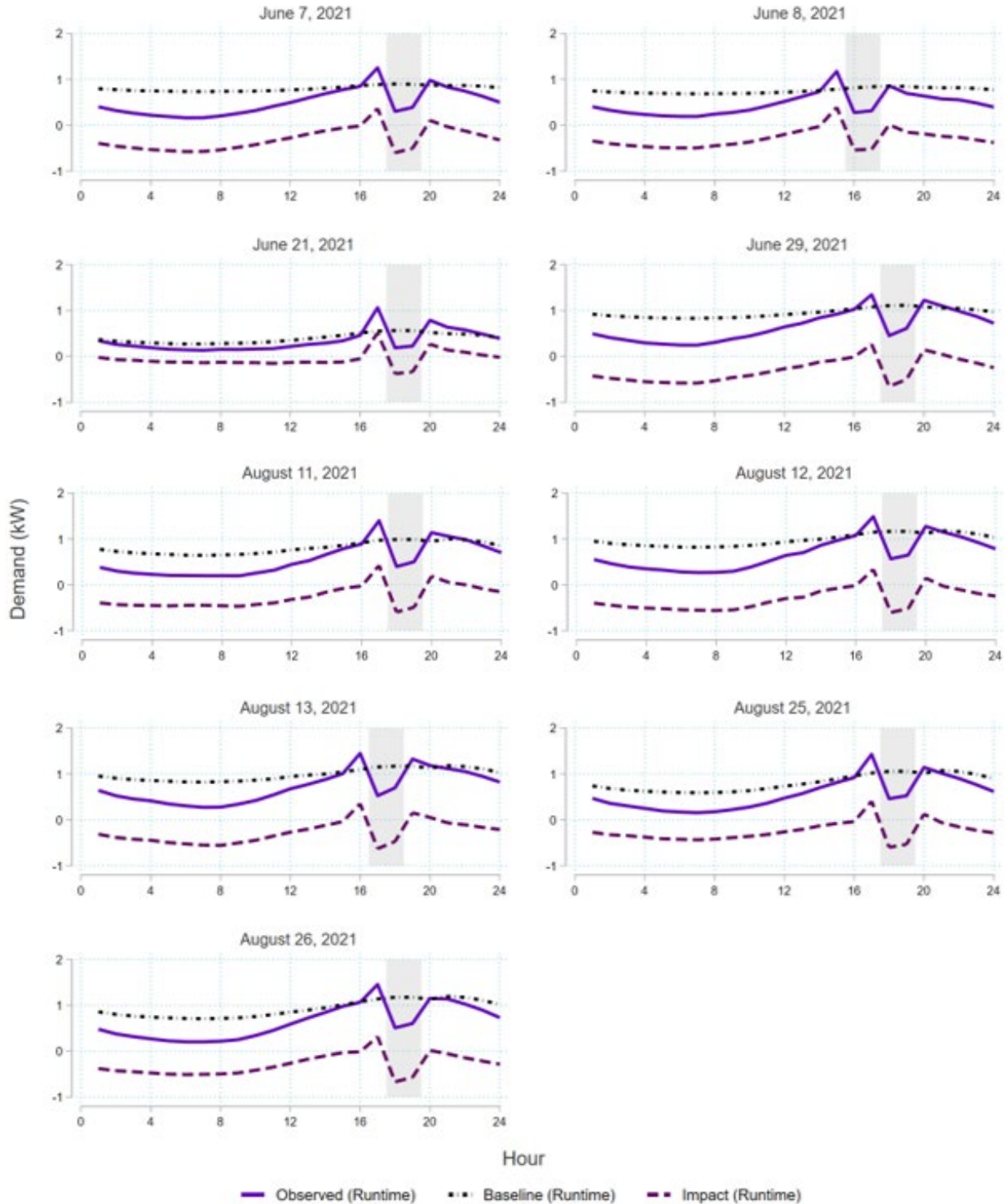
To understand these impacts at a program level, we use the average of the estimates produced during each event window and scale them by the number of thermostats targeted during the event. Across these four events, [Table 8](#) displays the average load shed from EnergyHub, the runtime, and the AMI. The AMI and runtime totals are less than the EnergyHub total because of the 3.5 kW connected load assumption that EnergyHub uses.

Table 8: Impacts of 2020 Residential Events

Date	Avg. Shed Per Thermostat (kW)			Aggregate Shed Across Targeted Thermostats (kW)		
	EnergyHub	Runtime	AMI	EnergyHub	Runtime	AMI
6/23/2020	0.78	0.46	0.47	628	377	383
7/9/2020	0.68	0.29	0.61	518	223	468
7/20/2020	0.97	0.60	0.57	2,100	1,299	1,226
8/11/2020	1.02	0.64	0.60	2,080	1,312	1,233
Average	0.86	0.50	0.56	1,332	803	827

The last set of events we can analyze are the nine events in the 2021 season. With only the runtime data available for this year, Figure 15 shows the demand reductions that we calculated using the 2.05 kW connected load assumption.

Figure 15: Residential 2021 Event Days



Each event depicts a clear drop in demand during the event hours, but Table 9 shows how these impacts are much lower than what is reported by EnergyHub, due to the differential in the connected load assumption. It is also important to note that the first two events of the 2021 DR season experience a much larger drop in aggregate kW, due to the dispatch of an additional 1,000 thermostats. These additional thermostats are only used to construct the baseline for the June 7 and June 8 event days and do not dilute the effect of the other event days.

Table 9: Impacts of 2021 Residential Events

Date	Avg. Shed Per Thermostat (kW)		Aggregate Shed Across Targeted Thermostats (kW)	
	EnergyHub	Runtime	EnergyHub	Runtime
6/7/2021	0.99	0.55	3,860	2,186
6/8/2021	0.99	0.53	3,760	2,064
6/21/2021	0.64	0.24	1,800	664
6/29/2021	1.02	0.43	2,850	1,194
8/11/2021	0.91	0.43	2,820	1,317
8/12/2021	0.94	0.45	2,900	1,382
8/13/2021	0.88	0.42	2,350	1,128
8/25/2021	0.93	0.46	2,870	1,416
8/26/2021	1.01	0.49	3,100	1,494
Average	0.92	0.44	2,923	1,427

The estimates that we produce across these three years and two datasets lead us to two conclusions:

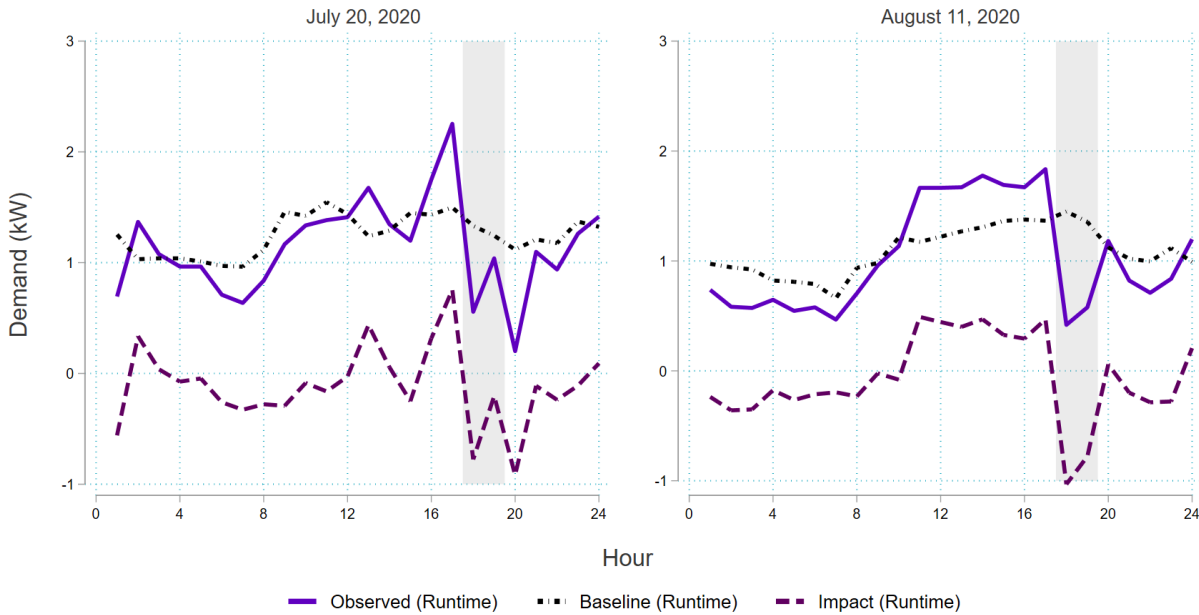
1. The most accurate way to calculate impacts on event days is to use AMI data. This is especially true of this program, due to the data limitations that the runtime data presents. The AMI data we received for the sample accounts was clean and almost all accounts had a full set of observations. Without missing observations from specific thermostat manufacturers, more accurate load shed estimates can be produced.
2. If AMI cannot be used to develop these impacts, the connected load assumption of 3.5 kW must be revisited. Using runtime data to produce impacts requires an accurate connected load assumption, since an overstated connected load will inevitably overstate the impacts of the program.

2.6.2 Small Business

The small business events were always dispatched on the same date and time as the residential events. These events were much smaller in scale, as the maximum number of targeted thermostats was 24. We only have runtime data for these customers, since we chose not to request AMI for this small sample of small business customers. Without both runtime and AMI data for a DR season, we are unable to create an adjusted connected load assumption and will use the 3.5 kW assumption for the 2020 and 2021 small business events.

Figure 16 shows the last two event days in the 2020 DR season. Due to OEM C devices being unable to be dispatched on the June 23 and July 9 event days, we only have runtime data for two thermostats, and therefore, cannot construct a valid baseline approximation for these two days.

Figure 16: Small Business 2020 Event Days



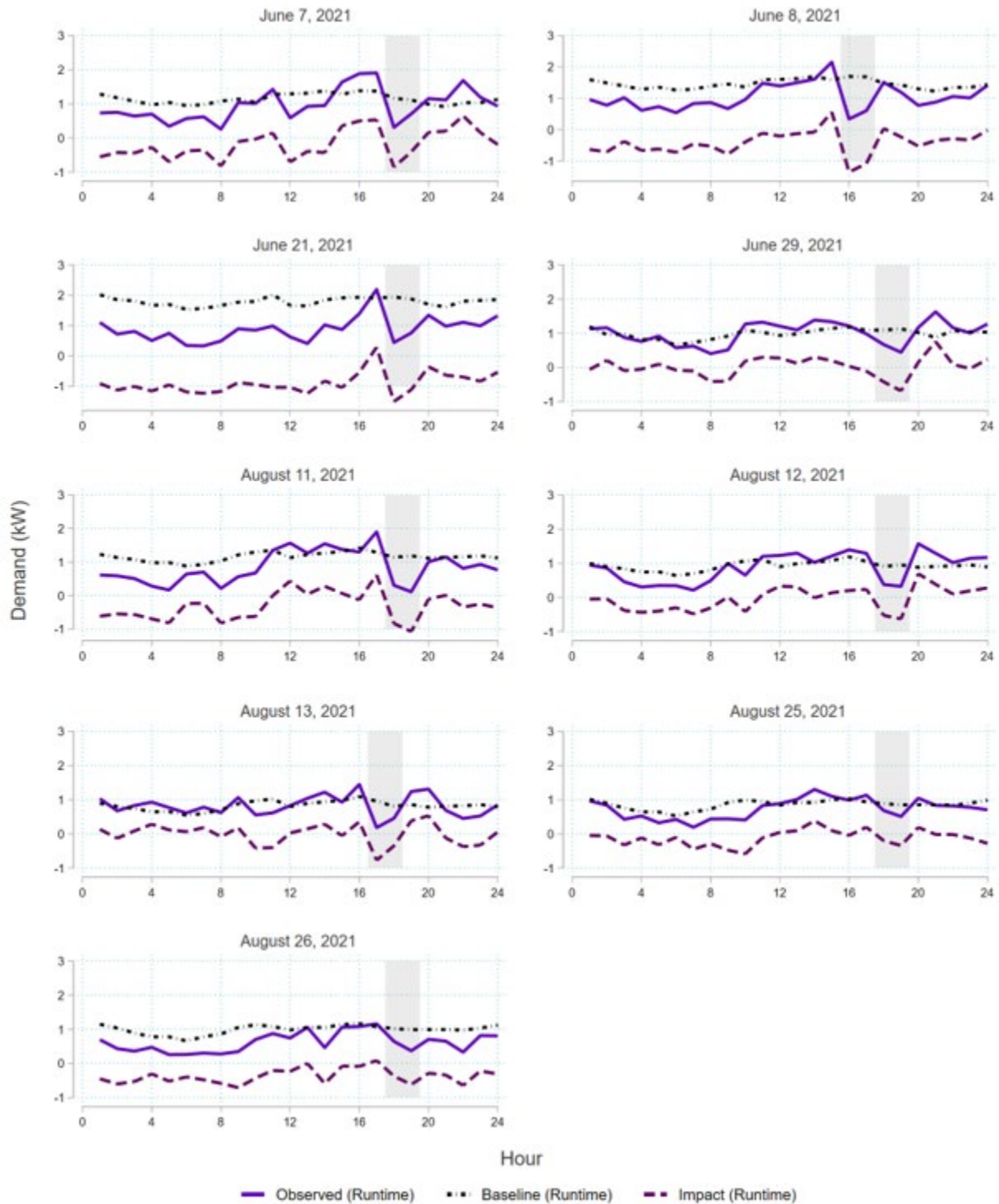
Due to the OEM C dispatch issues, we only report the aggregate load shed for the last two event days in the 2020 season in Table 10. The aggregate load shed for these two events days is constructed based on 11 devices. Due to this small sample size, the difference between the EnergyHub results and the evaluation results is likely not a statically meaningful difference.

Table 10: Impacts of 2020 Small Business Events

Date	Aggregate Shed Across Targeted Thermostats (kW)	
	EnergyHub	Runtime
6/23/2020	-	-
7/9/2020	-	-
7/20/2020	8.59	10.28
8/11/2020	16.31	20.81
Average	12.45	15.55

The 2021 events displayed in Figure 17 indicate a clear drop in demand during the event window. This is due to all thermostats being eligible for dispatch across all events.

Figure 17: Small Business 2021 Event Days



The first two events of the 2021 season were dispatched jointly with the residential customers, meaning we did not receive individual reporting for the small business program. Table 11 reports the aggregate shed for seven of the nine small business events, for which we received EnergyHub reporting.

Table 11: Impacts of 2021 Small Business Events

Date	Aggregate Shed Across Targeted Thermostats (kW)	
	EnergyHub	Runtime
6/7/2021	-	-
6/8/2021	-	-
6/21/2021	22.24	23.58
6/29/2021	23.60	10.01
8/11/2021	20.49	17.17
8/12/2021	7.64	7.51
8/13/2021	15.75	10.09
8/25/2021	12.54	6.47
8/26/2021	13.21	11.89
Average	16.50	12.39

Across the runtime data, the small business customers do reduce their demand during the event window, but since the program is small, savings are limited. Due to the small scale, most of the conclusions we draw from the residential customers can be similarly applied to the small business customers. Additionally, if this program were to be expanded, it is likely to produce larger and more accurate savings.

2.7 RECOMMENDATIONS

- **Thermostats should be required to be placed into a child group.** These groups are currently defined as “Residential Central AC” and “Small Business Central AC”. Allowing thermostats to sit without a child group creates an equity issue, since enrolled thermostats receive an incentive regardless of how often they were dispatched during a DR season. It also limits DR performance, as there is a subset of enrolled thermostats not dispatched during DR events.
- **Revise the current connected load assumption to 2.1 kW.** The connected load assumption of 3.5 kW appears to overstate the cooling load and ultimately the DR impacts of this program. Given the limitations that are presented when discussing the connected load calculations, this connected load assumption should be studied further.
- **Use AMI data where it is available.** There are many caveats when it comes to using the runtime data, including the need to calculate a more accurate connected load. Using the AMI data instead would allow for a more straightforward calculation of program kW impacts.

Section 3 United Illuminating Wi-Fi Enabled Heat Pump Water Heaters

Key Findings

- Heat pump water heater telemetry data is especially poor
 - EnergyHub claimed 0 kW in 2021
- HPWH are so efficient that there is limited opportunity for load shifting

3.1 PROGRAM OVERVIEW

This evaluation is based upon United Illuminating (UI) Company's Wi-Fi enabled HPWHs, which began in 2018. This program dispatches both summer and winter events. Generally, demand reductions during events tend to be low for three reasons:

1. HPWHs are highly efficient.
2. The number of enrolled customers is low.
3. The residential water heating end use is not as "load following" as air conditioning so there is limited diversified HPWH load during summer afternoon/evening Connecticut system peaks.

3.2 EVENT SUMMARY

Across the winters of 2019, 2020, and 2021, there were 11 demand response events called. [Table 12](#) details the small number of participants, and therefore, the small aggregate demand reductions during the events.

Table 12: Winter Event Day Characteristics

Event Date	HPWHs Targeted	Event Start	Event End	EnergyHub Aggregate Shed (kW)
1/11/2019	54	5 PM	7 PM	8.16
1/23/2019	54	5 PM	7 PM	2.68
2/28/2019	57	5 PM	7 PM	5.97
3/7/2019	57	5 PM	7 PM	4.43
1/6/2020	66	5 PM	7 PM	4.94
1/8/2020	66	5 PM	7 PM	0.60
1/22/2020	66	7 PM	9 PM	4.48

Event Date	HPWHs Targeted	Event Start	Event End	EnergyHub Aggregate Shed (kW)
1/30/2020	66	8 AM	10 AM	2.51
2/11/2020	66	7 PM	9 PM	5.55
12/16/2020	66	5 PM	7 PM	0.00
1/26/2021	66	5 PM	7 PM	0.00

In addition, across the summers of 2019, 2020, and 2021, there were 17 demand response events called. Like [Table 12](#) suggests, the demand reductions in [Table 13](#) are also small. EnergyHub reports zero impacts for all summer 2021 demand response events.

Table 13: Summer Event Day Characteristics

Event Date	HPWHs Targeted	Event Start	Event End	EnergyHub Aggregate Shed (kW)
7/17/2019	64	4 PM	6 PM	3.70
7/19/2019	64	5 PM	7 PM	3.12
7/30/2019*	64	4 PM	6 PM	2.20
8/8/2019	64	4 PM	6 PM	0.54
8/19/2019*	64	7 PM	9 PM	1.27
6/23/2020*	66	5 PM	7 PM	0.61
7/9/2020	66	5 PM	7 PM	3.09
7/20/2020	66	5 PM	7 PM	0.88
8/11/2020*	66	5 PM	7 PM	2.06
6/7/2021	66	5 PM	7 PM	0.02
6/8/2021	66	3 PM	5 PM	0.00
6/21/2021	66	5 PM	7 PM	0.00
8/11/2021	66	5 PM	7 PM	0.00
8/12/2021*	66	5 PM	7 PM	0.00
8/13/2021	66	5 PM	7 PM	0.00
8/25/2021	66	5 PM	7 PM	0.00
8/26/2021	66	5 PM	7 PM	0.00

3.3 DATA

3.3.1 Enrollment Data

HPWH participants are absent from the enrollment file.

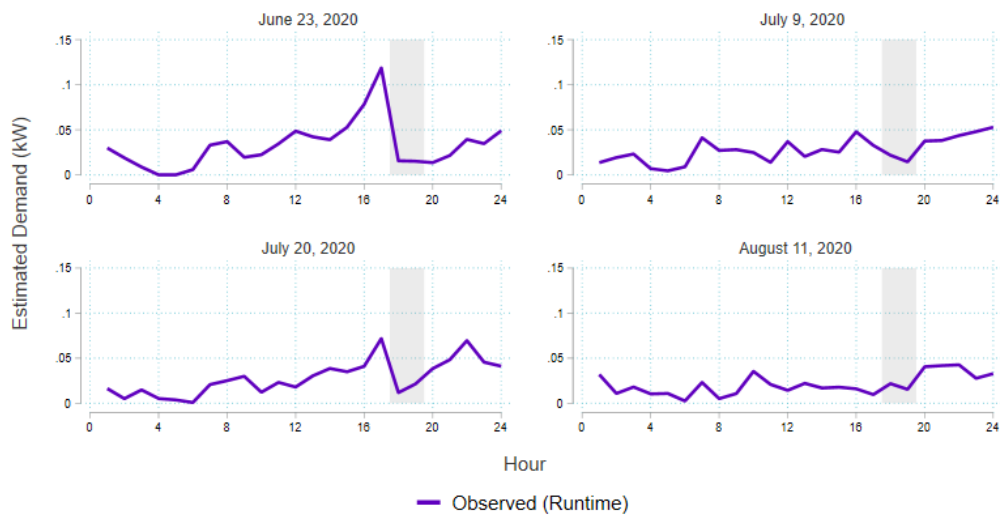
3.3.2 Runtime

The HPWH runtime file provides the estimated consumption during each 15-minute interval that the system is in operation. The only available file for HPWHs is the event-day runtime. Without the full summer runtime, we are unable to construct a baseline for the HPWHs. In discussions with EnergyHub, the study team learned that collection of reliable HPWH telemetry data from OEMs has been an ongoing challenge.

3.4 EX POST IMPACTS

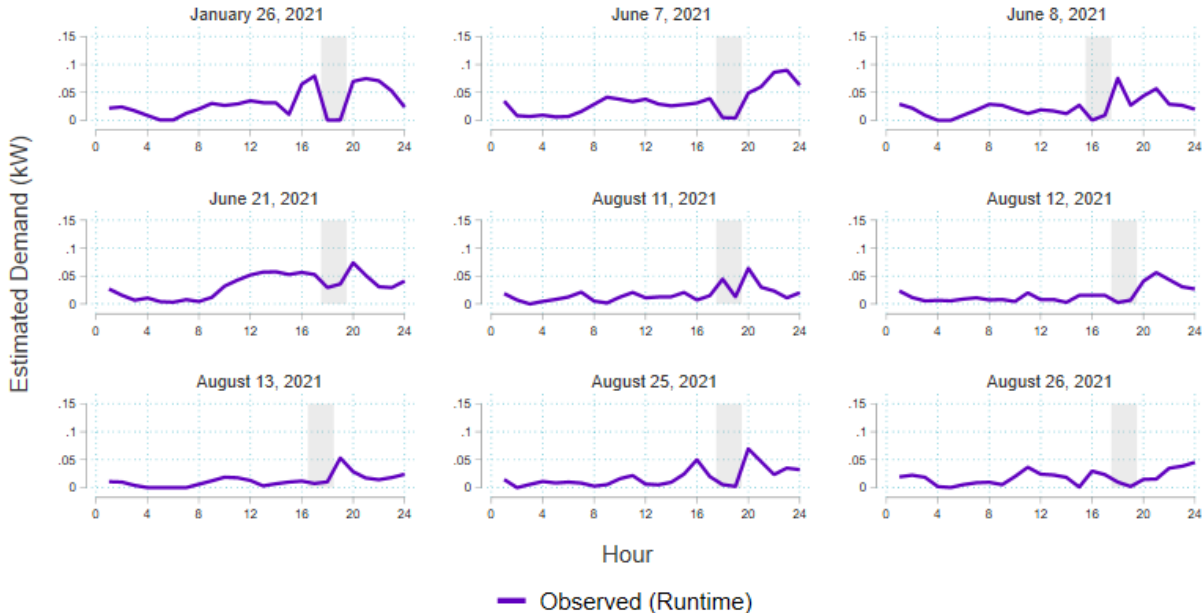
Without non-event day data, we were unable to construct baselines for event days. This ultimately means that impacts were not able to be calculated. Due to these limitations, [Figure 18](#) and [Figure 19](#) only plot the observed estimated demand from the runtime files.

Figure 18: HPWH 2020 Event Days



Although EnergyHub reports an average load shed of zero for all 2021 events, [Figure 19](#) appears to indicate that HPWHs were still dispatched during these events.

Figure 19: HPWH 2021 Event Days



Even without the ability to produce baselines, it is still evident that DR events for HPWH do not produce the same level of impacts that the thermostat program does. This is likely due to the high efficiency of these units, which limits the available load for available to actively manage. Without granular runtime data, we cannot determine the specific driver for small curtailable loads. Connecticut may be able to leverage planned HPWH metering in Massachusetts to better understand HPWH operating profiles across modes, customers, and through the day.²⁰

3.5 RECOMMENDATIONS

- **Target less efficient equipment.** The high efficiency of HPWHs severely limits the DR potential of the program. Even with the expansion of more devices, this program will be limited in kW impacts.
- **Target equipment with coincident loads.** Since HPWHs peak in the morning, targeting these devices will not adequately help reduce the peak system loads in Connecticut.
- **Do not enroll additional devices from manufacturers who cannot supply quality telemetry.** Reliable hourly or sub-hourly device-level operating data should be a prerequisite for inclusion in any connected device demand response program.
- **Leverage the storage capability of water heaters.** Tank water heaters can be used as a thermal storage device to store clean energy from solar during the day and reduce

²⁰ <https://ma-eeac.org/wp-content/uploads/Heat-Pump-Water-Heater-Impact-Study-Volume-1.pdf> discusses planned metering activities

ramping in the afternoon and evening.²¹ If pre-heating is used, it is important that there is a cold water mixing mechanism to prevent scalding.

²¹ https://www.dret-ca.com/wp-content/uploads/2021/04/pge_wattsaver_field_test_dret_final_report.pdf

Section 4 United Illuminating C&I Auto Demand Response Program

Key Findings

- Evaluated load reduction is substantially different than reported values.
- Getting accurate estimates of load reduction for a small number of sites with highly variable load will always be difficult.
- The standardized mean 10-of-10 with adjustment provides the best load reduction estimate.

4.1 PROGRAM OVERVIEW

This evaluation assesses United Illuminating (UI) Company's C&I Auto Demand Response (ADR) program. This program began as a pilot late in the summer of 2018. The pilot was designed to understand customer willingness to enroll in the ADR program, improve understanding of the DR automation process, and enhance distribution grid reliability through automation of event dispatch. The initial pilot targeted four to six large C&I customers serviced by the Ash Creek and Woodmont substations, which are predicted to require future load relief. Ultimately, three sites enrolled in the pilot with a total nominated load reduction of 266 kW. [Table 14](#) puts the pilot number in the context of program enrollment goals from 2020 through 2024. In fact, as of the summer of 2021, only seven sites were enrolled, with a total nominated load reduction of 481 kW. Based on communication with UI, we understand low enrollment is due at least in part to the additional challenges presented by the demand response automation process.

Table 14: Pilot Enrollment and Program Enrollment Goals

		Pilot	2020	2021	2022	2023	2024
Enrolled Sites	New	-	6	8	12	15	18
	Existing	3	3	9	17	29	44
	Total	3	9	17	29	44	62
Nominated kW	Annual Goal	-	528	704	1,056	1,320	1,584
	Total	266	794	1,498	2,554	3,874	5,458

To enroll, customers must commit a minimum of 50 kW of load reduction, have an operating energy management system (EMS), building management system (BMS), lighting control system (LCS), or programmable logic controller (PLC) that is connected to electric equipment that may be turned off during a DR event, and an interval meter. When a DR event is called, an event signal is sent from UI to the customer, through an OEM A gateway installed on-site, that initiates the predetermined control strategy to reduce load. The program has faced concerns regarding data

security due to the linkage between the demand response management system (DRMS) and the customer’s EMS/BMS. Addressing customer concerns can be difficult and cause time-consuming recruitment hurdles. UI is considering moving away from automated demand response and utilizing aggregators to increase program enrollment more quickly.

Up to 12 summer events and six winter events may be called annually, each event typically lasting two hours in duration but never more than four hours. Four two-hour events were called during the summer 2019 season of the pilot; however, no events were called during the winter 2019 – 2020, summer 2020, and winter 2020 – 2021 seasons. UI does not typically call winter events due to a 2018 Avoided Energy Supply Component (AESC) Study²² finding that winter demand response in New England is not cost-effective. Table 15 shows the date, maximum temperature, start and end time, and duration of the called events.

Table 15: Event Day Characteristics

Event Date	Max Daily Temperature (F)	Event Start	Event End	Event Duration (hours)
7/17/2019	91	4 PM	6 PM	2
7/19/2019	86	5 PM	7 PM	2
8/8/2019	86	4 PM	6 PM	2
8/19/2019	87	4 PM	6 PM	2

One objective of the program is to reduce the overall ISO-NE system peak. None of the called events included the 2019 ISO-NE annual peak installed capacity (ICAP) hour, which occurred July 30, 2019, from 5 to 6 PM, although the event periods generally overlap with ISO-NE seasonal peak hours. This represents a missed opportunity for participants to reduce their electric utility bills, though hitting the ICAP can be difficult as the system peak is not known in advance and hence is dependent on correctly forecasting it

Pilot incentives included a monthly reservation payment and a performance-based payment, while the program only offers performance-based incentives. Performance or DR event load reduction is determined using 15-minute interval data. The difference between customer baseline load and actual load during DR events equals load reduction. Incentive payments are based on the average hourly DR event load reduction. This program uses the ISO-NE mean 10-of-10 baseline methodology, with an additive same day load adjustment to determine customer baseline load. For sites with solar generation, the methodology was modified to discard baseline days with substantial solar PV system output and the adjustment was not always utilized due to solar PV output. The program’s baseline methodology modifications for sites with solar generation were not defined further, and the evaluation team relied on observing the shared performance calculations to replicate the reported claims. In these instances, the baseline methodology appears to have been modified on an arbitrary basis.

Because the program only had three participating sites in 2019, this evaluation could look more closely at individual site results to examine the quality of individual load reduction estimates.

²² [AESC-2018-17-080.pdf \(synapse-energy.com\)](https://www.synapse-energy.com/AESC-2018-17-080.pdf)

4.2 METHODS

4.2.1 Data

To support the evaluation of the ADR pilot program, the evaluation team received event performance data and AMI data for all 2019 and 2020 pilot participants. A spreadsheet was received for each participating site. Each spreadsheet included the site address, substation, summer and winter season load reduction nominations and strategies, and event performance calculations. The AMI data provides consumption for each participating site for the 2019 and 2020 summer DR seasons. Consumption is provided in two separate data streams, an import stream, and an export stream. The import stream represents energy consumed by a site from the grid. The export stream represents energy that is exported from the site to the grid, for sites with behind-the-meter generation. Together, the two streams represent the site's 'net' energy use with respect to the grid. Importantly, these data do not capture behind-the-meter generation output that is not exported to the grid. Collectively, this data was used to verify reported event performance and perform the regression analysis. As no DR events were called in 2020, the analysis focuses on the 2019 summer DR season.

4.2.2 Baselines

A fundamental difference between load reduction and generation as a resource is that it is not possible to directly observe load reduction. Instead, measurement of load reduction necessarily means comparing observed load to a counterfactual load that would have occurred in the absence of dispatch. This counterfactual, or baseline, load is the estimate had the demand response event not occurred. The difference between the estimated baseline load and observed load during an event is the load reduction. Load reduction is positive if the observed load is less than the baseline and negative if the observed load is greater. Both the calculated baseline load and calculated load reduction are estimates and necessarily subject to some estimation error. In the discussion that follows, we review the baseline methodologies used in this evaluation to estimate load reduction, load characteristics that affect estimation accuracy, and the relationship between program design, baseline choice, and accuracy. The evaluation team estimated impacts by comparing observed load on event days with three different baselines (unadjusted mean 10-of-10, symmetrically adjusted mean 10-of-10, and regression). The ADR program generally uses a symmetrically adjusted 10-of-10 baseline as the basis for settlement. These load reduction estimates are reported by the initiative implementer. The symmetrically adjusted 10-of-10 baseline is the basis for the evaluation team's estimate of load reduction. Each baseline's calculation method, advantages, and disadvantages are described below.

4.2.2.1 Unadjusted Mean 10-of-10

This baseline collects load data for the ten most recent eligible non-event days preceding the event day; eligible days include non-holiday weekdays. The unadjusted baseline shape is calculated as the average load for each interval (e.g., 15-minute) of the day over those ten days. When there are missing load data for the ten most recent days, additional lookback days may be used to collect ten days with sufficient data to estimate baseline load.

The unadjusted 10-of-10 baseline is effective for customers with consistent, non-weather-sensitive load, has minimal potential for manipulation, and allows for preparatory decreases in

load (e.g., shift cancellation) without affecting baseline load or load reduction estimates. However, the unadjusted 10-of-10 is inflexible for customers with weather-sensitive load or load variability unrelated to weather; this baseline can substantially understate baseline load and load reduction for customers with weather sensitivity if events are called on extreme weather days. The ADR program calls events to mitigate load during peak hours which is correlated with extreme weather.

4.2.2.2 Symmetrically Adjusted Mean 10-of-10

The symmetrically adjusted baseline shifts the unadjusted 10-of-10 shape to meet observed load during pre-event hours (adjustment window). For an additive adjustment, the magnitude of the shift for each interval of the day equals the difference between observed load and the unadjusted baseline load during the adjustment window. Symmetric adjustments may shift the unadjusted baseline upward or downward. However, the resulting baseline may not be less than zero, except for customers where behind-the-meter generation is present.

To the extent that load during the adjustment window is representative of non-event day load, a symmetrically adjusted baseline generally improves the accuracy and reduces bias of baseline load estimates for customers with weather-sensitive load or load variability unrelated to weather. The ADR program provides day-ahead event notification; the adjustment window is the 15-minute interval occurring immediately prior to the event period. Though the use of automated DR may address this concern, generally, the use of day-ahead event notification increases the potential that load during the adjustment window will be influenced by the event. Event-related effects during the adjustment period include pre-cooling and baseline manipulation. Pre-cooling can inflate the adjusted baseline for weather-sensitive customers. Deliberate ramp-ups may be difficult to distinguish from preparatory load changes. Short term notification occurring after the baseline adjustment period reduces the potential for baseline manipulation and variability due to preparatory load changes. Since the adjustment window for the ADR program occurs immediately prior to the event period, the adjustment window would necessarily need to be moved back to implement same-day event notification. As a point of reference, ISO-NE uses this baseline with short-term notification and an adjustment window defined by the three five-minute intervals occurring prior to the notification. This baseline, with a 1-hour adjustment window beginning 2 hours prior to the event period, is also the basis for the evaluated load reduction impacts of program administrators' C&I active demand response initiatives in MA, CT, and NH.²³ Since day-ahead notification is used for these active demand response initiatives, event-related effects on the baseline continue to be studied.

4.2.2.3 Regression

Ex post regression-based baseline methods offer the promise of using a full set of seasonal load data and weather data to characterize a site's load on event days. The goal is to explain as much of the natural load variation at the site through a combination of variables that address varying schedules as well as weather sensitivity. By capturing the underlying site-level load dynamics, regressions provide an estimate of expected load during a DR event, which offers a day-specific alternative to settlement baselines that take advantage of same-day load adjustment. The intent

²³ [Report Title \(ma-eeac.org\)](http://ma-eeac.org)

is that regression approaches offer a reasonable, unbiased baseline unaffected by strategic load altering behavior on the event day and other days.

For this analysis, five regression specifications are estimated for each site across multiple degree-day bases. The optimal model is chosen using adjusted R-square. All specifications (defined in the Regression Specifications appendix) are combinations of schedule variables (hour, weekday, and monthly dummies) and weather variables. Estimating a variety of specifications facilitates identifying a model that best characterizes the unique load dynamic at each site. As is always the case with load modelling, the information available to support the independent variables in the model is limited. These models are limited to different combinations of schedule variables and temperature, linear and squared. For this work, we did not include two relatively common approaches. The Lawrence Berkeley National Laboratory's (LBNL) time-of-week-temperature (TOWT) model has a singular, segmented linear weather trends that is not ideally suited for forecasting load on the most extreme days of the year. The other common specification estimates afternoon hour load as a function of load earlier in the day, a form of lagged dependent variable approach. This approach is not included because the intent of the regression approach is to offer an alternative to a same day adjusted baseline.

The regression can be an effective baseline for customers with either stable or weather-sensitive load and can control for weather without a day-of-event adjustment. Because the ex post regression analysis does not have a day-of-event adjustment, it has the potential to facilitate analysis of changes in consumption that occur before and after the event. However, because the regression summarizes all data from the summer, it can systematically underestimate load reduction for customers with unscheduled or unreported shutdowns and is erratic for customers with high load variability unrelated to weather (e.g., operational variability).²⁴

4.3 IMPACT VERIFICATION

This section verifies the reported load reduction estimates and baseline used as the basis for settlement. This is the estimate of load reduction that determines incentive payments for an enrolled asset. The ADR program generally uses a mean 10-of-10 baseline with a symmetric, additive same day load adjustment as the settlement baseline. For sites with solar generation, the settlement baseline methodology is reportedly modified to discard baseline days with substantial solar PV output and to selectively not apply the baseline adjustment where there is solar PV output during the adjustment window on the event day. The program's settlement baseline methodology modifications for sites with solar generation were not defined further, and the evaluation team relied on observing the shared performance calculations to replicate the reported claims as a rule was unable to be inferred. For the 2019 summer DR season, two of the three participating sites have solar generation present. To verify the reported results, the evaluation team replicated the program settlement baseline methodology. Table 16 shows the total nominated, reported, and verified load reduction estimates for each of the four summer 2019

²⁴ Granderson, J, M Sharma, E Crowe, D Jump, S Fernandes, S Touzani, and D Johnson 2021. "Assessment of Model-Based Peak Electric Consumption Prediction for Commercial Buildings" https://eta-publications.lbl.gov/sites/default/files/crowe_-_assessment_of_model-based_peak_.pdf

DR events and the average across the season. The verified load reduction estimates are the evaluation team’s attempt to replicate the reported load reduction estimates.

Table 16: Summer 2019 Performance Verification

Event Date	Load Reduction (kW)		
	Nominated	Reported	Verified
7/17/2019		368.5	254.8
7/19/2019	266.4	271.1	191.6
8/8/2019		511.4	454.1
8/19/2019		561.7	488.0
Average	266.4	428.2	347.1

The evaluation team was able to replicate the reported event load reduction within 1 kW for two of three sites. For the other site, verified load reduction estimates were roughly one-tenth the size of the reported results. The evaluation team confirmed with UI that the original performance calculations were in error. Table 17 shows the nominated, reported, and verified load reduction estimates for this site, asset 1, for each of the four summer DR events and on average across the season.

Table 17: Summer 2019 Performance Verification - Asset 1

Event Date	Load Reduction (kW)		
	Nominated	Reported	Verified
7/17/2019		126.7	12.7
7/19/2019	67.4	88.3	8.3
8/8/2019		63.2	6.3
8/19/2019		83.4	8.4
Average	67.4	90.4	8.9

Table 18 compares the evaluation team’s load reduction estimates used to verify reported performance with the evaluation team’s load reduction estimates using a mean 10-of-10 baseline with a symmetric, additive, same day load adjustment across all sites and aligned with standard practices. The results that use the adjusted baseline methodology are shown in the Evaluated Adjusted Settlement Baseline column.

Table 18: Summer 2019 Performance – Verified vs Evaluated

Event Date	Load Reduction (kW)		
	Nominated	Verified	Evaluated Adjusted Settlement Baseline
7/17/2019		254.8	639.0
7/19/2019		191.6	237.8
8/8/2019	266.4	454.1	137.8
8/19/2019		488.0	(152.6)
Average	266.4	347.1	215.5

There are several reasons for the differences between verified load reduction estimates and load reduction estimates constructed using the evaluation team's standard symmetrically adjusted mean 10-of-10 baseline, including,

- For sites with solar generation, the settlement baseline methodology was reported to have been modified to discard baseline days with substantial solar PV system output and to selectively not apply the baseline adjustment where there is solar PV output during the adjustment window on the event day. However, the selection of baseline days and application of the same day load adjustment are not well defined and appear to have been done on an arbitrary basis. The evaluation team's adjusted settlement baseline results use the ten most recent eligible baseline days prior to the event day and consistently apply the same-day load adjustment.
- In one instance, less than ten days were used to construct the settlement baseline. The evaluation team always uses ten days to construct the settlement baseline.
- In another instance, an event day was used to construct the settlement baseline. The evaluation team never uses ineligible baseline days (e.g., event day, weekend, holiday) to construct the settlement baseline.
- For one site, baseline days were discarded due to an unanticipated significant decrease in load around the July 4 holiday, including days that were eligible for a DR event to be called. In this instance, additional lookback days were used instead of the days with decreased load. The evaluation team always uses the ten most recent eligible baseline days prior to the event to construct the settlement baseline.
- The settlement baseline methodology adjustment window specifies the 15-minute interval occurring immediately prior to the event as the adjustment window. The evaluation team specifies the adjustment window as the hour beginning two hours prior to the event period.

For the remainder of the evaluation, the mean 10-of-10 baselines used to estimate load reduction are constructed using this consistent methodology.

4.4 IMPACT EVALUATION

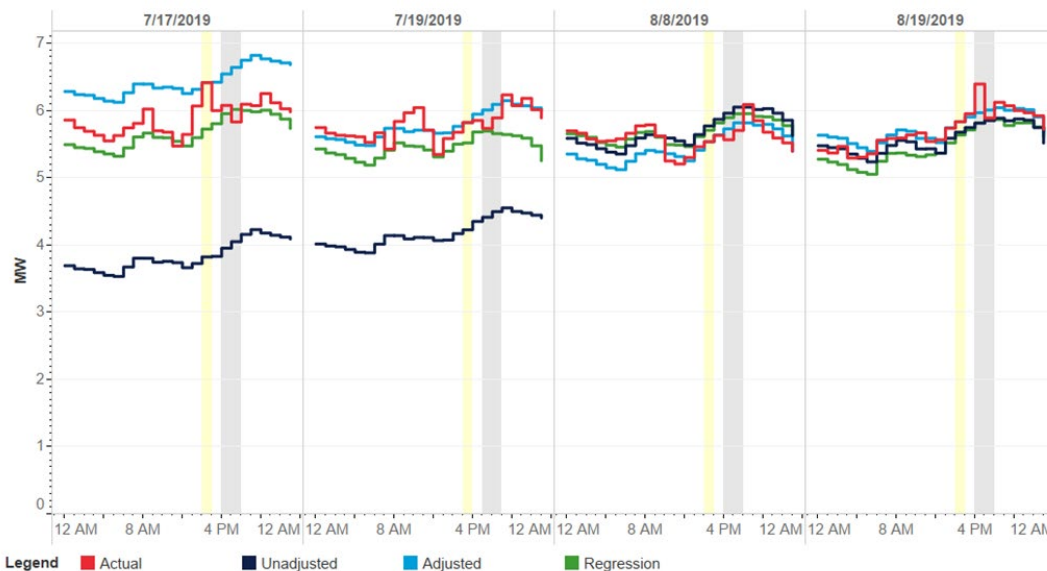
This section compares the evaluation team’s load reduction estimates using an unadjusted mean 10-of-10 baseline, an adjusted mean 10-of-10 baseline, and a regression baseline. Table 19 shows the total nominated, unadjusted, adjusted, and regression baseline results for each of the four summer 2019 DR events and the average across the season. No DR event was called for the ICAP hour, which occurred July 30, 2019. Further discussion of these results follows this table.

Table 19: Summer 2019 Evaluated Performance

Event Date	Load Reduction (kW)			
	Nominated	Evaluated Unadjusted Settlement Baseline	Evaluated Adjusted Settlement Baseline	Regression
7/17/2019	266.4	(1,952.2)	639.0	31.0
7/19/2019		(1,357.2)	237.8	(136.7)
8/8/2019		371.0	137.8	287.6
8/19/2019		(308.6)	(152.6)	(290.3)
Average	266.4	(811.7)	215.5	(27.1)

Figure 20 provides a visual representation of each of the four summer 2019 DR event days for all sites in aggregate. The figure shows actual load (red line), the unadjusted 10-of-10 baseline (dark blue line), the adjusted 10-of-10 baseline (light blue line), and the regression baseline (green line). The adjustment period is highlighted in yellow, and the event period is shaded in grey.

Figure 20 : Summer 2019 Event Days



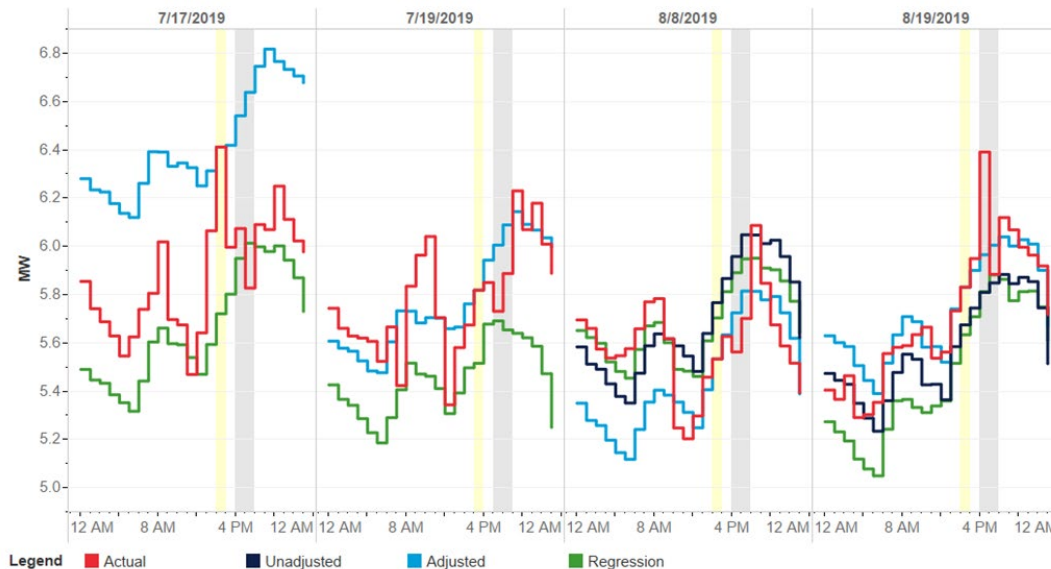
These program level results are primarily driven by asset 3, the largest participating site. This site’s load is typically between 3.5 and 6 MW and has the largest nominated load reduction of 154 kW. Assets 1 and 2 have on-site solar generation and typical peak loads in the range of 100 kW and 600 kW, respectively. Due to substantial solar generation, asset 2 is generally in the process

of ramping back into positive consumption at the onset of event hours as solar generation diminishes. In combination assets 1 and 2 represent approximately 10% of the overall program peak load.

The unexpectedly low unadjusted settlement baseline for the first two events is driven by asset 3. Site load indicates that this site reduced activity for more than a week around July 4. Those partial shutdown days occurred on days when a demand response event could have been called, were unreported, and are thus included in the settlement baseline day choice. As a result, the unadjusted settlement baseline for asset 3 is roughly 30% lower than pre-event load on those event days and has an additive adjustment that represents 50% of baseline load on the first event day. Oddly enough, the substantial apparent overestimate of load reduction for the July 17 event is not caused by asset 3 but by extreme adjustments of the other smaller assets, likely driven by variability in solar generation.

Figure 21 is the same as Figure 20, but without the unadjusted baseline on July 17 and July 19, due to this baseline skewing the perspective of the graph. This makes it easier to distinguish between the actual load and the baselines. The next three sections examine each of the sites individually. Each site demonstrates various challenges of estimating load reduction in the DR program context.

Figure 21: Summer 2019 Event Days – Zoomed In



4.4.1 Asset 1

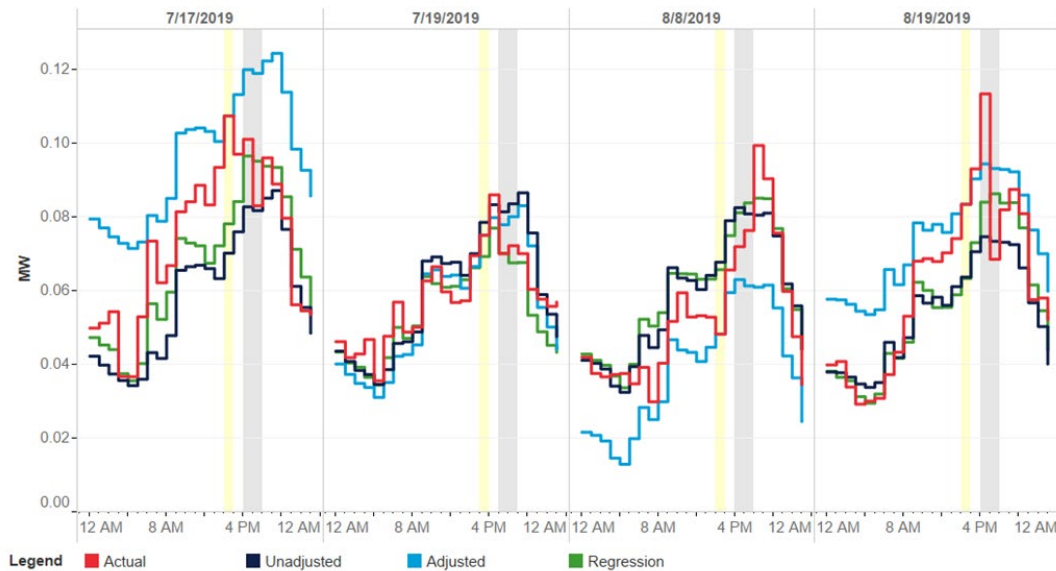
Asset 1 is the site with the smallest evident peak load at or around 100 kW on event days. This site is a library on the Ash Creek substation. Summer nominated load reduction is estimated at 67.4 kW and comes from setpoint adjustment on 7 RTUs. If the nomination is correct, the load reduction would represent a roughly 70% reduction in the evident peak load for the site. Load reduction representing 70% of consumption is an unrealistic expectation of the impact provided by setpoint adjustments, however this site also has a 65-kW solar array on its roof. If the observed load is net load incorporating generation of that magnitude, then midday consumption at the

library could still be more than 50% greater than the net load data available. Looking across the summer, the variation across days provides plenty of evidence that solar is present but the effective generation appears to be closer to half of the array’s rating.²⁵

Figure 22 provides the asset-level plots of the four event days. The plots demonstrate both day to day and hour to hour variability. Adjustment period load varies from above 100 kW to below 50 kW across the events reflecting both an overall reduction and later shift of load on lower days that would be consistent with varying solar generation. Daytime load may also reflect weather sensitivity as general load levels track with temperature across the event days. During the day, the regression baseline, with its built-in weather adjustment, varies as expected relative to the unadjusted baseline with its recent weather focus. On the first and last event days, which are the hotter two days, the regression baseline shows higher early evening load than the unadjusted baseline. In both cases, though, the regression does not appear to fully account for the higher load correlated with higher temperature in the late afternoon.

It is also worth noting that moving the adjustment period back to a more standard distance from the start of the event substantially degrades the consistency and average magnitude of load reductions. A 15-minute adjustment period immediately prior to an event is too easy to game or be unintentionally affected by pre-cooling. Such an adjustment period will always reduce the variation of load reduction because of that proximity but it does not necessarily produce a more valid estimate of load reduction.

Figure 22: Asset 1 Summer 2019 Event Days



In addition to the day-to-day variation, there are numerous hourly movements of over 10 kW that occur at different hours across the event days. This variability is generally illustrated by the variation of the actual load shape relative to the regression baseline. The day to day and hourly

²⁵ Asset 2’s substantial solar array provides a clearer indication of the variation in solar generation due to cloud cover, etc. There are numerous days when generation is less than 25% of typical sunny days.

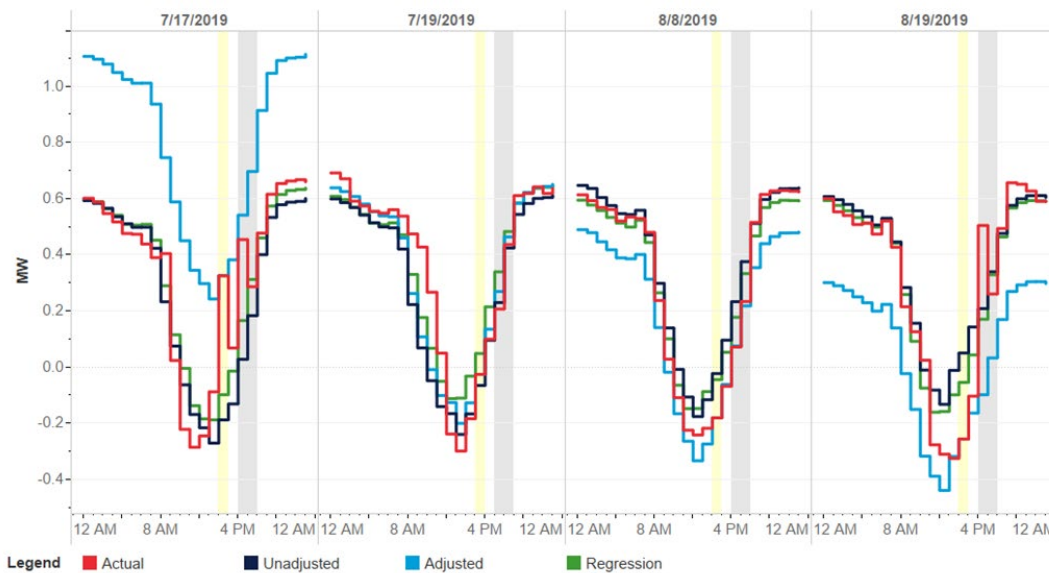
variability are illustrated more dramatically by the adjusted baseline during the July 17 event. The combination of particularly high load during the single adjustment hour and an unadjusted baseline shape with a much later peak produces an obviously exaggerated adjusted baseline during the event period. The unadjusted baseline reflects a shape with more solar than occurred on the event day.

4.4.2 Asset 2

Asset 2 is on the Ash Creek substation and has site-level overnight load of between 600 and 700 kW with a few days above 800 kW. The site is a water treatment facility with a substantial solar array. While we do not know the rated capacity of the solar array, the load shape is consistent with a relatively flat industrial load that has a solar array capable of generating 800 or more kW in the middle of the day. Summer ADR Program nominated load reduction is estimated at 45 kW and comes from setpoint adjustment on 7 RTUs. The load reduction represents a roughly 7% reduction of peak load for the site.

Figure 23 provides the event day load shapes. The event period between 4 and 6 PM occurs when solar generation is decreasing. The nominated load reduction is small compared to the typical hourly increase in load during that part of the day. Without the solar array, or with full site energy consumption reconstituted with array generation, this site would likely be a good candidate for the demand response. Solar generation is an unknown, variable driver of the net load shape. Controlling for the solar generation with either a settlement or regression baseline is unrealistic.

Figure 23: Asset 2 Summer 2019 Event Days



The August 19 event appears to provide an example of the inherent challenge of combining solar generation and DR. Both assets 1 and 2, both with solar and located in the same town have substantial spikes in load during the first hour of the event when we would expect to see load reduction. A possible explanation for this coordinated spike is that some form of cloud cover moved in during that hour that reduced solar generation considerably. Asset 2 increases

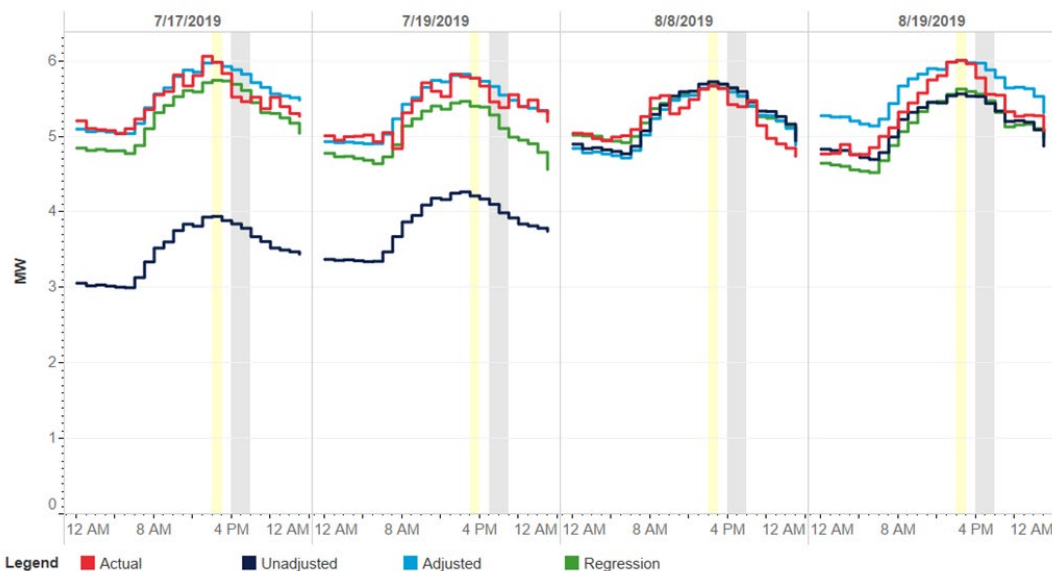
consumption by 600 kW between the hour before the event and the first hour of the event. Similar variability in the solar generation could also explain the unrealistic adjustment that is applied for the July 7 event. While, in general, the net load shape for this asset appears relatively smooth, this masks a substantial amount of variability that undermines the predictability of load during an event. The regression results indicate that substantial variation is present. The adjusted R-square of the best regression model is only .88 compared to .96 and .99 for the other assets' best models.

4.4.3 Asset 3

Asset 3 is a large production facility with a typical peak load approaching 6 MW on the Woodmont substation. The nominated load reduction is 154 kW and is produced by raising setpoints on eight RTUs. The load reduction represents a 2.5% reduction in load at the site. Ceteris paribus, smaller changes in load are more difficult to detect than larger ones. As unexplained variability in load increases, so does the minimum detectable load reduction.

Figure 24 provides the event day load shapes for Asset 3, the site that shut down for over a week after the July 4 holiday. This is evident in the low unadjusted settlement baselines for the first two events. Despite the substantial adjustments for the first two events, the adjusted settlement baseline appears to provide a reasonable baseline across the first three events. The last event, August 19, has a decidedly steeper ramp than any of the baselines and the mismatch of the adjusted baseline outside of the event hour is stark. In general, the similarity of all load shapes reflects the consistent, low variability load at this site. There is some shifting that the regression and unadjusted baselines are unable to track.²⁶ The regression for this asset explains over 99% of variation and still performs very poorly on two of the four events.

Figure 24: Asset 3 Summer 2019 Event Days

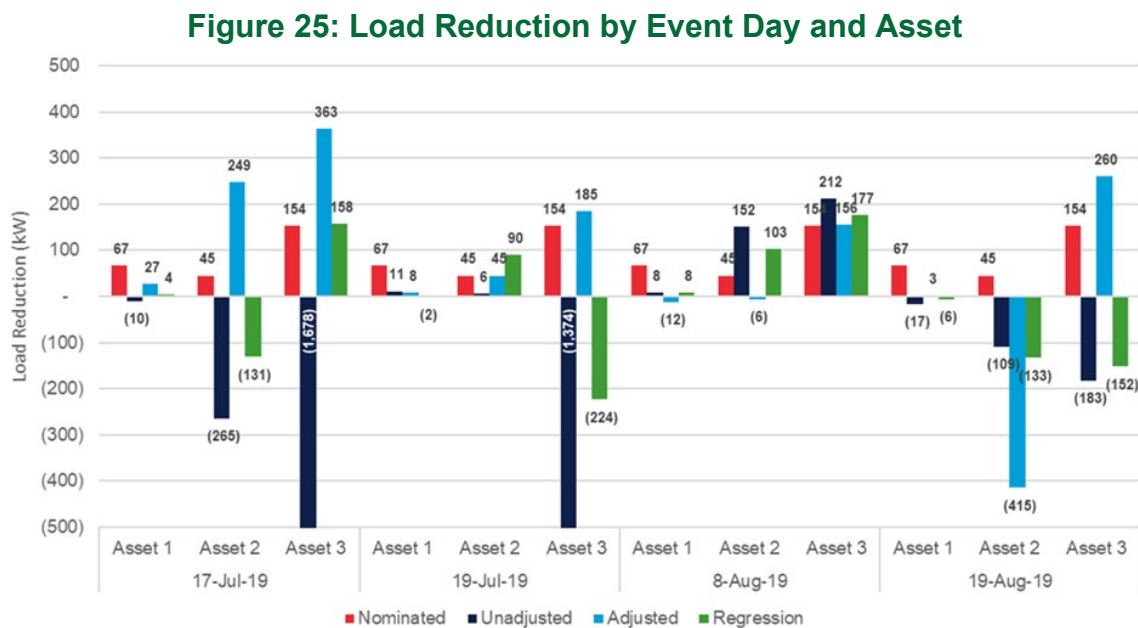


²⁶ The shutdown period for asset 3 was pulled out of the regression or the regression baseline would be much closer to the unadjusted baseline.

4.5 FINDINGS AND RECOMMENDATIONS

This section summarizes the findings and provides considerations for the summer 2019 season of UI's Automated Demand Response program.

Figure 25 summarizes individual asset performance across baselines by event. Nominated load reduction is included to provide a way to distinguish between the expected magnitude of load reduction at the sites. When looking at this graph, it is important to remember that the nominated load reduction for asset 1 appears unrealistic given the typical load shape (4.4.1). In general, the nominated load reduction may not be any closer to actual load reduction than any of the baseline estimates.



The key findings are:

- None of the baselines provide consistent estimates of load reduction across assets and events. The regression baseline generally performs similar to or better than the unadjusted settlement baseline but does not track event day load well. The adjusted settlement baseline more frequently provides reasonable estimates of load reduction but also produces some clear overestimates (due to flatter shapes adjusted up to steeper, descending load) and one extreme negative (most likely related to the solar generation falling). Of these baselines, the adjusted settlement baseline provides the best load reduction estimates.
- Asset 1 has a nominated load reduction that appears unrealistic given its typical load shape. A more realistic nomination would be one tenth the magnitude.
- Asset 2 has a substantial solar array that outpaces site consumption on sunny summer days.
- Asset 3 has the lowest variability of the three assets, the smallest load reduction on a percentage basis, and is the only asset to produce positive load reduction estimates for

all events. The first and last events appear to be over-estimates due to a beneficial adjustment and overly flat baseline shape, respectively.

- This kind of variation of load reduction estimates at the customer level is not unusual but is typically obscured in the total load reduction of a program with more participants. Given the number of participants evaluated, their unique load characteristics, and the nature of load reduction, the best ex ante estimate of summer load reduction is the average event load reduction during the summer 2019 using the adjusted settlement baseline. The average event load reduction during the summer 2019 using the evaluated adjusted settlement baseline is 216 kW.
- Sites with solar generation and net load shapes may not be suitable for DR programs using baselines for settlement. If the full site-level consumption can be re-constituted, then baseline settlement is possible. If not, an unknown and variable driver of load levels has the potential to render any baseline ineffective.
- Site-level regression as an ex post evaluation technique for demand response is risky. As discussed in the paper referenced in section 4.2.2.3, the problem of regression models underestimating load in a DR context is recognized in the industry.

The key recommendations for the program are:

- If the program intends to use baselines for settlement, it needs to implement a clear settlement baseline methodology consistently applied and well-documented to increase fairness and transparency, particularly for the settlement baseline methodology for sites with solar generation, which was not well-defined nor systematically replicable.
- Same-day notification can improve baseline adjustments by decreasing the probability of event-related load effects or gaming during the adjustment window. If participants are notified of an event prior to the interval(s) used to adjust the baseline, there is a chance of influencing loads (up or down) by knowledge of and preparation for the upcoming event. Since the demand response strategies are automated, we believe the potential benefit of a more accurate baseline greatly outweighs the minimal risk of additional program burden and/or dissatisfaction on participating customers.
- The program should modify the adjustment period if load reduction strategies involve pre-event changes to load (e.g., pre-cooling). For example, Massachusetts C&I demand response programs use the hour beginning 2 hours prior to the event as the adjustment window. This adjustment window decreases the likelihood that pre-event changes to load are included in the baseline adjustment. Further, the longer adjustment period can also help to decrease the impact of a spurious change in load leading to a poor baseline adjustment. ISO-NE provides same-day event notification and uses the 15-minute interval preceding event notification as the adjustment window. If the program changes to a same-day event notification, this is also another possibility.
- The program should establish a load predictability requirement. Successful assessment of load reduction for a DR program, either for settlement or ex post evaluation, requires a balance of site-level load predictability with the expected magnitude of load reduction. A

relatively simple baseline analysis can determine the accuracy of a baseline for event-type days for comparison to nominated load reduction.

- The presence of solar is a substantial issue for the measurement of DR. For sites with solar generation, the program should obtain the solar output data to reconstitute site load and then estimate customer baseline load and load reduction.
- The number of pilot participants and events evaluated is an issue for the measurement of DR. The program should build on the pilot's level of customer participation and event frequency to create a grid resource that is more reliable and more evaluable than the limited 2019 pilot activity. Greater numbers of participants can be expected to provide more reliable load reduction and load reduction estimates, whereas smaller population sizes, like in this study, are prone to the nuances of one or two customers having an outsized influence on the overall result. Addressing the additional challenges presented by the demand response automation process, as noted earlier, may naturally increase enrollment rates and the number of participants.

Section 5 Eversource Connected Solutions Wi-Fi Air Conditioners

Key Findings

- Telemetry data quality varies by manufacturer
- Window AC and mini-splits have small diversified peak loads.
- The program reduced the loads by 38%, as compared to the 58% reduction from the Eversource Wi-Fi thermostat program.

5.1 PROGRAM OVERVIEW

This section summarizes the evaluation of Eversource Company's Connected Solutions Wi-Fi Air Conditioner demand response program. This program was launched in the summer of 2020 and continued through the summer of 2021. It is a direct load control program comprised of residential customers who enroll their own room air conditioner and ductless mini-split devices. These devices are demand response enabled from the factory and do not require any supplemental control equipment. It provides a demand response participation option from homes unable to participate in the Wi-Fi thermostat program due to a lack of central A/C.

Participants receive \$20 per season in exchange for allowing Eversource to make slight adjustments to the setpoint or cycle the unit on and off up to six days between June and September. The program did not include pre-cooling of homes prior to events during summer 2021. ThinkEco, the implementation contractor used by Eversource in 2021, works with manufacturers of Wi-Fi enabled equipment. The window AC and mini-split OEMs each have their own proprietary algorithms. The study team was unable to get specific details from ThinkEco about the control strategies implemented during summer 2021. Using "state change" telemetry from the enrolled connected devices from summer 2021, we quantify the load reductions of these events using non-event day consumption patterns to estimate what customer loads would have been absent demand response dispatch. The telemetry data did not include information about opt-outs so the study team could not examine the frequency which participants overrode the programmatic adjustments.

5.2 EVENT SUMMARY

During the 2021 season, six events were called. Each event lasted three hours, beginning at 4 PM and ending at 7 PM. [Table 20](#) details the maximum temperature during the event²⁷, as well as the event start and end time for each event. Additionally, those days that ISO-NE identified as

²⁷ Based on the Hartford Brainard Field weather station.

the peak day for each month are denoted with an asterisk. Eversource dispatched demand response (DR) events on one of the three ISO-NE peak days.

Table 20: Event Day Characteristics

Event Date	Max Event Temperature (F)	Event Start (EDT)	Event End (EDT)
7/15/2021	88	4 PM	7 PM
7/27/2021	87	4 PM	7 PM
8/6/2021	86	4 PM	7 PM
8/12/2021*	90	4 PM	7 PM
8/13/2021	94	4 PM	7 PM
8/26/2021	91	4 PM	7 PM

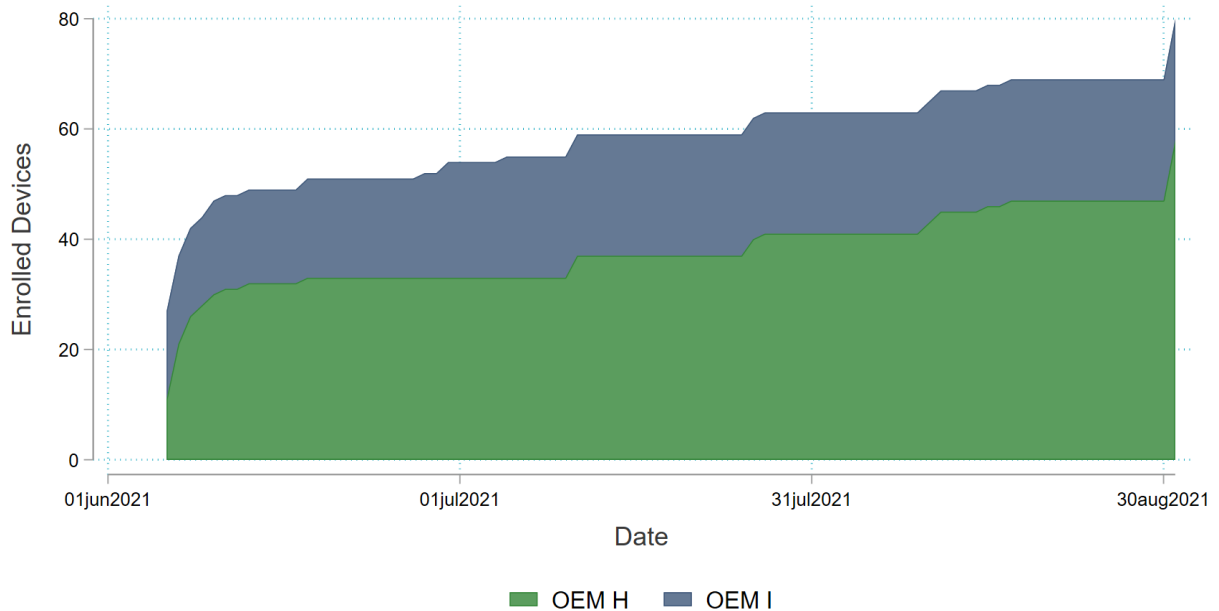
* ISO-NE System Load Peak Day

5.3 DATA

5.3.1 Enrollment

To understand how participation grew throughout the 2021 DR season, we use the first state-change observation as a proxy enrollment date. ThinkEco does not have access to device telemetry until a customer enrolls. [Figure 26](#) shows the number of enrolled devices by manufacturer during this time. This graph inevitably leaves out 11 OEM H devices that were reported in the enrollment file but were not present in the state-change data as well as an unknown number of OEM J devices, since we cannot construct a proxy enrollment date.

Figure 26: Enrolled Devices by Day



On the first event date, July 15, 2021, there were 59 devices with state-change observations. By the sixth and final event on August 26, 2021, enrollment grew by an additional ten devices. Across the entirety of the season, OEM H devices constitute most of the program.

5.3.2 State Change

The data we received was in a state-change format. This data type creates a new observation each time a setting on the device changes, which includes:

- **Target temperature.** An increase/decrease in the desired temperature of the space.
- **Power status.** A switch between power “On” and “Off.”
- **Operation mode.** A change across “Cooling,” “Heating, and “Saving” modes.

This data is unlike interval runtime or AMI data because it does not provide a metric to measure the device’s electric demand. Ultimately, we must convert this state-change data to interval data, as detailed in [Section 5.4.1](#), to estimate the electric demand.

Understanding the granularity of this manufacturer data is important in the development of the program since it is necessary in evaluating performance. Without data that can accurately be converted to electric demand, program performance remains unknown. [Table 21](#) details the current data quality issues of the three manufacturers in the program.

Table 21: Data Quality Issues

Manufacturer	Devices in State-Change	Data Quality Issues
OEM H	47	<ul style="list-style-type: none"> 11 devices are present in the enrollment file but not in the state-change file. OEM H indicated these were legacy models which they could not retrieve state change data for Multiple state-change records on a single minute <ul style="list-style-type: none"> Can be corrected by receiving the state-change with millisecond timestamps
OEM I	22	<ul style="list-style-type: none"> Did not provide an enrollment file Does not report ambient temperature
OEM J	--	<ul style="list-style-type: none"> Did not provide any data

Based on these issues, the evaluation team recommends Eversource set minimum viable data quality standards prior to adding new manufacturers or enrolling additional devices from OEM J and OEM I.

5.4 METHODS

5.4.1 Converting to Electric Demand

To convert to electric demand, we must assume within the state-change data. Since the OEM H and the OEM I data came in different formats, the assumptions we made differed slightly. For OEM H data, we defined the unit as drawing power when the power status was “On,” and the ambient temperature was higher than the target temperature. The OEM I data does not record an ambient temperature, so we assume the unit is drawing power when the power status is “On.”

OEM H and OEM I each provided the model numbers that correspond with the devices in the state-change data. Using these model numbers, the evaluation team pulled the cooling capacity in BTU per hour and the CEER (window AC) or EER (mini splits). The power draw in Watts is calculated by dividing the cooling capacity by the CEER or EER. We then convert this into kW by scaling the calculation by 1,000. Table 22 details the calculated power draw by manufacturer, as well as overall. The average power draw of the mini split units was slightly higher (0.97 kW) than the window AC units (0.94 kW).

Table 22: Power Conversion

Device Type	Mean (kW)	Minimum (kW)	Maximum(kW)
OEM H	0.91	0.62	1.43
OEM I	0.97	0.50	1.59
Total	0.95	0.50	1.59

To take this power conversion and produce interval data, we determine how many minutes the unit is drawing power within a 15-minute interval. This is then converted to a variable detailing the

percent of minutes the unit is running, by dividing the runtime by 15. This percentage is then multiplied by the calculated power conversion to get the electric demand in that 15-minute period.

$$\text{Electric Demand} = \frac{\text{Minutes Running}}{15} * \text{Power Conversion}$$

5.4.2 Baseline

To estimate demand reductions, it is necessary to estimate what electric load would have been on an event day in the absence of a DR event. This estimated load, commonly referred to as a counterfactual, can be approximated using a baseline. We rely on non-event day demand patterns to estimate our baseline. Using these non-event days allows us to model how customers would have consumed energy on an event day had they not been dispatched. Baselines are simply a tool to produce demand reductions estimates, not the end itself. They help filter out noise and explain variation, to allow the demand reduction to be more easily detected.

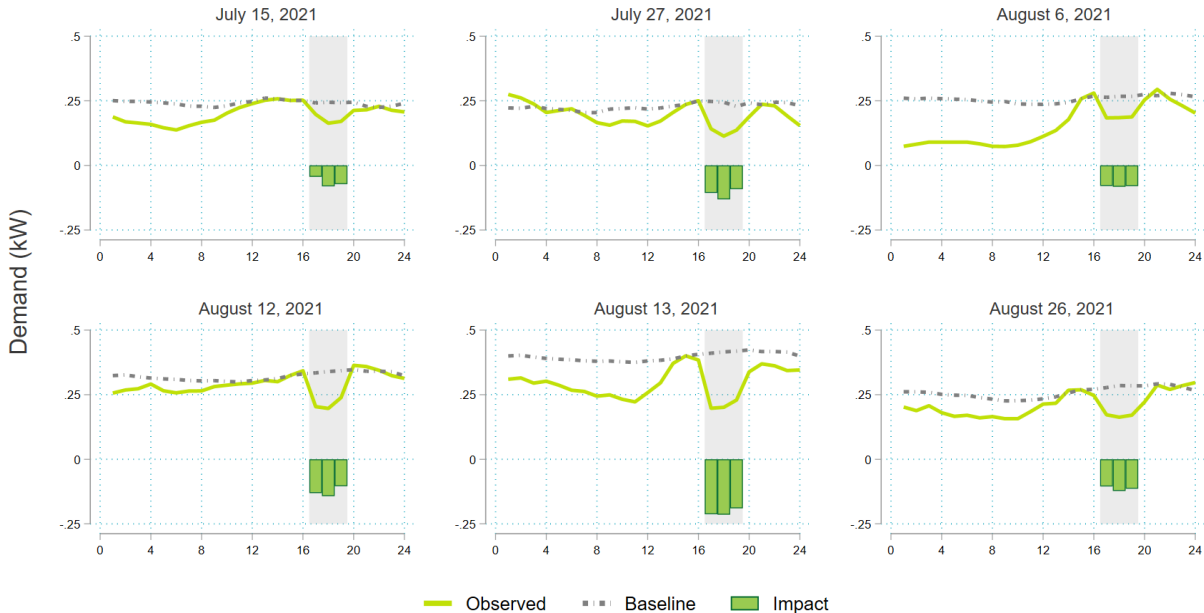
The baseline we use to construct impacts is the ISO-NE settlement baseline. This is a mean 10-of-10 baseline with an additive adjustment, which is based upon an hour buffer and an hour pre-event adjustment interval. The last ten eligible days are defined as non-holiday, non-event weekdays.

The change in energy use, or the impact, is computed by taking the difference between the baseline and the observed demand during the event. We calculate this impact on a per-device basis. To construct an aggregate load shed value, these per-device impacts are averaged across the three-hour event window and then scaled by the number of enrolled devices, plus the additional 11 devices that do not have state-change.

5.5 EX POST IMPACTS

For the six events in 2021, we were able to analyze both the OEM H and OEM I device demand reductions. The per-device impacts are displayed in [Figure 27](#). Each event day displays a clear drop in demand during the set of event hours, but average cooling load does not drop to zero – meaning that units are still operating during event hours, just at a reduced level.

Figure 27: ConnectedSolutions 2021 Event Days



To understand these impacts at a program level, we use the average of the impacts produced during each event window and scale them by the number of devices enrolled on that event day. Table 23 displays these results for OEM H and OEM I devices, as well as the overall program results.²⁸

Table 23: Impacts of 2021 ConnectedSolutions Events

Date	Max Event Temp (F)	Avg. Shed Per Device (kW)			Aggregate Shed Across Enrolled Devices (kW)		
		OEM H	OEM I	Overall	OEM H	OEM I	Overall
7/15/2021	88	0.07	0.06	0.07	3.22	1.40	4.61
7/27/2021	87	0.11	0.10	0.11	5.93	2.26	8.12
8/6/2021	86	0.06	0.12	0.08	2.94	2.67	5.96
8/12/2021	90	0.09	0.19	0.13	4.98	4.28	9.81
8/13/2021	94	0.22	0.18	0.21	12.13	4.07	16.03
8/26/2021	91	0.09	0.15	0.11	5.44	3.32	9.10
Average	89	0.11	0.14	0.12	5.77	3.00	8.94

These impacts are small, in part due to the small reference loads. The highest reference load of 0.42 kW occurs on the August 13, 2021, event day and the average reference load across the six

²⁸ Given the data issues encountered across the programs, and the associated potential for systematic uncertainty, the study team believes the statistical uncertainty of the impacts likely overstates the precision of the findings and elected not to show in this report.

events was 0.32 kW. With a connected load of 0.95 kW and an average reference load of 0.32 kW, the implied coincidence factor from the study results is 34%.²⁹ Given the data received, it is difficult to determine if the assumptions we make converting the state-change to the interval data are artificially decreasing the program loads, or if the program loads are relatively small to begin.³⁰ Regardless of the answer to this question, on average, this program delivers a 38% reduction³¹ in cooling load during events hours.

5.6 RECOMMENDATIONS

1. **Data quality should be assessed prior to device manufacturers being included in the program.** The absence of quality telemetry creates uncertainty regarding the program performance. If a device manufacturer is unable to provide this data, there is no assurance that demand response events are being dispatched to the customer. If the manufacturer can provide it, there is no way to quantify their load reductions.
2. **Device manufacturers should provide guidance on when a unit is drawing power.** Since state-change data only reports when a device changes operation mode, there is no clear definition on when a device is drawing power or for how many minutes the device is running. Even with adequate state-change data, assumptions must still be made about when the unit is drawing power and how much power it is drawing. Without clarification from the manufacturers, the assumption implemented could either be over-stating or under-stating the load of these devices.
3. **Eversource should work with manufacturers to make their curtailment algorithms more aggressive.** The economics of this program are likely to be challenging with load impacts of 0.12 kW per device. There is nothing the program can do about equipment size and reference load, but participating devices could be curtailed more completely. The study estimates a 38% reduction in cooling demand during events, on average. For comparison, the Eversource Wi-Fi Thermostat percent reductions were estimated at 58% during summer 2019.³² In [Section 2](#) of this report, the study estimates an average percent reduction in cooling demand of 46% for UI smart thermostats during demand response events.

²⁹ The 2022 Connecticut PSD list a summer coincidence factor of 23.2% for ductless mini-split heat pumps and does not provide a coincidence factor for Window AC units.

³⁰ The study team is not aware of other studies that document the conversion logic for telemetry from these OEMs. The vendor (ThinkEco) is no longer in business. Ideally, if the program continues under a new vendor it would be good for the vendor to work out those details with the OEMs.

³¹ $0.12 / 0.32 = 38\%$. 0.32 kW is the average baseline, 0.12 kW is the load reduction, and 0.95 kW is the connected load. The implied coincidence factor is thus $0.32 / 0.95 = 34\%$. Because this analysis uses runtime data the calculation actually goes the other way (e.g., the devices are on 34% of the hour and have a 0.95 kW connected load, therefore the baseline is 0.32 kW),

³² <https://ma-eeac.org/wp-content/uploads/2019-Residential-Wi-Fi-Thermostat-DLC-Evaluation-Report-2020-04-01-with-Infographic.pdf> Table 4-6

Section 6 Value of DR as a Grid Resource

Key Findings

- DRR have been recognized for the ability to decrease load on the system, but DRR now play an expanded role in energy markets.
- DRR also provide non-system benefits that are not easily quantifiable, but which can be monetize to facilitate policy goals.
- Participation in the ISO-NE FCM could open additional revenue streams for DR programs, but the risks of performance penalties should be weighed against the potential benefits.

6.1 INTRODUCTION

This section examines electric demand response programs (DR) in Connecticut offered by The United Illuminating Company (UI) and Eversource Energy to assess future opportunities for program expansion and enhanced value creation. The evaluators researched and reviewed electric DR programs offered in other jurisdictions including New England states—Maine, Massachusetts, New Hampshire, Rhode Island, Vermont—California, Hawaii, and New York. Recent efforts across the country to reduce the negative impact of fossil fuel generation have created opportunities to align policy with program design to create synergies that achieve both climate and energy goals. The value of demand response is driven by policy objectives, recognition of its benefits, and the ability to quantify and account for those benefits in the evaluation of the programs.

All benefits and value of DR are predicated on the effective deployment of it as a resource when and where needed. Program designs have features that can directly affect the ability of a program to achieve its capability (signed up resource levels) vs its actual reduction. Program specific evaluations often provide the true peak impact achieved and can, as appropriate, provide derating factors to ensure the curtailment of a specific DR program is properly accounted for in the forecasting and planning that rely on them

Cost-benefit assessment criteria that limit the recognized benefits of demand response programs also limit the value that demand response programs can claim. The section reviews the benefits recognized in the current programs offered by UI and Eversource and highlights potential untapped value streams in current and future programs. This research also shows that program administrators in other states evaluate DR programs using a wider range of benefits and costs, including NEIs, that produce synergies between climate and energy policy goals while boosting the cost-effectiveness of energy programs including Demand Response.

The primary objectives of this section are to:

- Research and document the system benefits from the demand response programs administered by UI and Eversource in Connecticut,
- Identify additional benefits that can be claimed by existing and future programs, and
- Provide insight into the potential options for expansion of DR in Connecticut.

The section begins with a review of the demand response programs and pilots offered by UI and Eversource in Connecticut, including the cost-effectiveness method and associated benefits captured by the programs. In addition, we examined programs offered in other states, program objectives, and benefits streams considered by each program's benefit cost analysis. In reviewing the programs, the evaluators identified untapped value streams available to the current program offerings as compared to programs in other jurisdictions. The evaluators then highlighted opportunities to expand the program offerings. Interviews were conducted with UI implementation staff to gather an understanding of their rationale for the design choices and any changes to their intent from the original plans.

6.2 OVERVIEW OF DEMAND RESPONSE

Utilities have the responsibility of maintaining system reliability through load management. Advancing technologies has allowed customers to become real-time DR resources with an accompanying ability to actively participate in utility load management programs triggered by electricity price signals or other forms of dispatch notification. Engaging and deploying these capabilities captures value related to generation, deferred utility transmission and distribution infrastructure investments, , enhanced system reliability, and electric energy price suppression. Independent System Organizations/Regional Transmission Organizations (ISOs/RTOs) have also provided an expanded opportunity for utilities to aggregate and control smaller DR resources to meet system reliability needs and that of customers while opening additional value streams. ISO/RTO coordination recognizes the importance of DR to managing loads and increasing regional market efficiency through the mechanism of Forward Capacity Markets (FCM).

In addition, technological advancements—such as advanced metering infrastructure (AMI) and connected devices such as smart thermostats and smart inverters—have allowed for the increased dissemination of individual and grouped distributed energy resources—including DR. Program administrators are better able to aggregate, communicate with, control, and measure the impact of devices when a DR event is called. State governments have focused on energy conservation to manage costs in the face of increasing energy costs and complexity in the electricity markets. The State and Local Energy Efficiency Action Network described demand response as typically being a more cost-effective way to meet demand than an increase in supply—at least in the short term.³³ These programs incentivize targeted load reduction during limited peak hours to capture system benefits for all rate payers.

³³ State Approaches to Demand Reduction Induced Price Effects: Examining How Energy Efficiency Can Lower Prices for All: Industrial Energy Efficiency & Combined Heat and Power Work Group. State & Local Energy Efficiency Action Network. December 2015.

The forward-looking nature of the organized wholesale electricity market is influenced by load reduction from utility programs—including programs administered by aggregators (i.e. Curtailment Service Providers). The current evaluation methodology of UI and Eversource demand response programs recognize the system benefits of these load reductions. Avoided costs as based on AESC 2021 and incorporated into the utility benefit-cost test include values for generation capacity, avoided transmission and distribution (both PTF and local), Demand Reduction Induced Price Effect (DRIPE), and reliability. The demand resources offered by customers fall into two categories: active demand resources, which requires active participation by the individual customers, and passive demand resources, which occurs through energy efficiency upgrades or distributed energy resources.³⁴ Although DR resources in UI and Eversource do not actively participate in the ISO-New England (ISO-NE) Forward Capacity Market, the measures in these programs “still provide indirect system benefits by impacting ISO New England’s forecast of load, which is one of the inputs used to develop prices in the capacity market.”³⁵

Neither Eversource nor UI currently participate in Active Demand Capacity Resource (ADCR) in the ISO-NE market. ISO-NE provides a base payment for availability and a pay for performance payment when they call on resources to reduce load, in response to a supply constraint condition. ISO-NE also imposes penalties for non-performance or under performance. These penalties would need to be weighed against the potential benefits of participation. Working with qualified Curtailment Service Providers or aggregators can help to mitigate some of this risk while providing the technical expertise to support ISO-NE participation. Since ISO-NE events are called to respond to supply constraints, while the utility DR program events are called to mitigate peak demand conditions, there is expected to be, and has historically been, overlap between these two programs.

Since the implementation of Price-Responsive Demand (PRD) by ISO-NE in 2018, most “PRD resources primarily served as capacity and operating reserve resources available for dispatch at very high offer price.”³⁶ However these resources have seen low dispatch levels averaging 5.5³⁷ MW and 3.2³⁸ MW in the day-ahead market in 2019 and 2020, respectively. This has totaled an energy market revenue of \$1.5 million and \$1 million in respective years. The estimated total capacity payment over the two years was \$70 million, averaging over 400 MW of capacity supply obligation in each year.^{39,40} PRD resources now include:

³⁴ Demand Response in ISO-NE Market and Planning. Connecticut Department of Energy and Environmental Protection. October 27, 2016.

³⁵ Avoided Energy Supply Components in New England: 2021 Report. Prepared for AESC 2021 Study Group. Released March 15, 2021.

³⁶ 2019 Annual Market Report. ISO New England Inc. Internal Market Monitor. June 9, 2020.

³⁷ Id at 108.

³⁸ 2020 Annual Market Report. ISO New England Inc. Internal Market Monitor. June 9, 2021.

³⁹ 2019 Annual Market Report. ISO New England Inc. Internal Market Monitor. June 9, 2020.

⁴⁰ 2020 Annual Market Report. ISO New England Inc. Internal Market Monitor. June 9, 2021.

- Active demand capacity resources (ADCRs) – resources that can participate in the Forward Capacity Market (FCM). ADCRs will replace the resource type known as real-time demand-response resources.⁴¹
- Demand response resources (DRRs) – resources that can participate in the energy and reserve markets.⁴²
- Demand response assets (DRAs) – the physical entities delivering megawatts which cannot directly participate in the marketplace.⁴³

ISO NE allows individual DRAs less than 5 MWs within the same dispatch and reserve zone to be aggregated to form a DRR of greater than 5 MWs.⁴⁴

6.3 ELECTRIC DEMAND RESPONSE IN CONNECTICUT

Demand response (“price responsive load”) in the state of Connecticut has played a role in the energy market for almost two decades. In the 2001 Annual Legislative Report, market impact on energy prices created an opportunity for increased load management. The role of the distribution companies’ programs was highlighted as important to “facilitating this new and beneficial type of interruptible-load supply (ILS).”⁴⁵

Since 2001, the reported demand savings from conservation and load management programs have increased from approximately 63MW of peak reduction in 2000⁴⁶ to 104 MW of peak reduction in 2020—Active Demand response accounted for 65.5 MW and approximately 17,600 customers participated in Active Demand response.⁴⁷ The demand response resources can be aggregated to participate in organized wholesale markets by offering services in energy, capacity, or ancillary service markets.⁴⁸ In bulk electricity markets across the country, demand resources compete to provide supply in the energy services. ISO-NE “supports resource providers who wish to bid DR resources into its market.”⁴⁹ If a DR participant bid clears the market, “the DR participant is paid for its supply at this clearing price.” Also, in the real-time market the DR participant can receive additional compensation for additional supply or may be required to pay for any shortfall in delivery at the real-time price.⁵⁰ We do note from staff interview that UI and Eversource are not capturing revenue from DR programs in the FCM due to the risks associated with non-

⁴¹ Price-responsive Demand. ISO New England Inc. <https://www.iso-ne.com/markets-operations/markets/forward-capacity-market/fcm-participation-guide/price-responsive-demand>

⁴² Id.

⁴³ Id.

⁴⁴ Price Responsive Demand Overview. ISO New England Inc. June 2017. <https://www.iso-ne.com/static-assets/documents/2017/11/20171107-webinar-prd-overview.pdf>

⁴⁵ Report of the Energy Conservation Management Board Year 2000-2001 Program and Operations. January 31, 2001.

⁴⁶ Id at 11.

⁴⁷ Energy Efficiency Board 2020 Programs and Operations report. March 1, 2021.

⁴⁸ Demand Response in ISO-NE Market and Planning. Connecticut Department of Energy and Environmental Protection. October 27, 2016.

⁴⁹ Demand Response Compensation Methodologies: Case Study for Mexico. National Renewable Energy Laboratory. June 2018.

⁵⁰ Id at 15.

performance or underperformance penalties.⁵¹ Customers in Connecticut benefit, indirectly, from energy and capacity price suppression stemming from avoided investments in system. Additional value streams driven by active participation of DR resources in the FCM—such as direct compensation for reducing system load during demand response events, and capacity payments for the nominated capacity declared dispatchable during a curtailment event are not captured in UI and Eversource programs as seen in other programs around the country.

6.4 BENEFITS AND COST OF DEMAND RESPONSE

The benefits of DR programs stem from several value streams that occur to customers, the electric distribution systems, and wholesale markets. GTM Research reports that “value streams can be found at all levels of the grid.”⁵² Demand response value streams occur to the customers through system peak mitigation that results in avoided demand and energy charges. Value streams that accrue to distribution utilities include avoided distribution line losses, enhanced reliability, and T&D infrastructure investment deferral. Wholesale market value streams can also include black start (i.e., capable of starting without an outside electrical supply), spin/non-spin reserves, energy, and capacity.⁵³ Table 24 shows a listing of the universe of benefits of demand response to customers, utilities, and transmission system operators according to research performed for the US congress.⁵⁴

⁵¹ Based on interview with utility staff.

⁵² Demand Response Value Streams Have Yet to Be Fully Monetized. Greentech Media. May 8, 2017.

⁵³ Demand Response Value Streams Have Yet to Be Fully Monetized. Greentech Media. May 8, 2017.

⁵⁴ Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005.

Table 24: Benefits of Demand Response

Type of Benefit	Recipient(s)	Benefit		Description/ Source
Direct benefits	Customers undertaking demand response actions	Financial benefits		<ul style="list-style-type: none"> • Bill savings • Incentive payments (incentive-based demand response)
		Reliability benefits		<ul style="list-style-type: none"> • Reduced exposure to forced outages • Opportunity to assist in reducing risk of system outages
Collateral benefits	Some or all consumers	Market impacts	Short-term	<ul style="list-style-type: none"> • Cost-effectively reduced marginal costs/prices during events • Cascading impacts on short-term capacity requirements and LSE contract prices
			Long-term	<ul style="list-style-type: none"> • Avoided (or deferred) capacity costs • Avoided (or deferred) T&D infrastructure upgrades • Reduced need for market interventions (e.g., price caps) through restrained market power
		Reliability benefits		<ul style="list-style-type: none"> • Reduced likelihood and consequences of forced outages • Diversified resources available to maintain system reliability
Other benefits	<ul style="list-style-type: none"> • Some or all consumers • ISO/RTO • LSE 	More robust retail markets		<ul style="list-style-type: none"> • Market-based options provide opportunities for innovation in competitive retail markets
		Improved choice		<ul style="list-style-type: none"> • Customers and LSE can choose desired degree of hedging • Options for customers to manage their electricity costs, even where retail competition is prohibited
		Market performance benefits		<ul style="list-style-type: none"> • Elastic demand reduces capacity for market power • Prospective demand response deters market power
		Possible environmental benefits		<ul style="list-style-type: none"> • Reduced emissions in systems with high-polluting peaking plants
		Energy independence/security		<ul style="list-style-type: none"> • Local resources within states or regions reduce dependence on outside supply

Inherently, DR programs require investments in control equipment by the participants and utilities with accompanying responsiveness from the customers to reduce load and capture the intended value. Automation of demand response is enabled by technologies such as smart meters and control equipment installed at the premise. To generate sufficient participation, program administrators develop programs with the appropriate level of incentives and educational campaigns to raise customer awareness of programs. Program support in the form of offsetting control equipment installation costs, in addition to the DR curtailment performance incentive, can increase participation in demand response programs but also increase costs that will need to be considered in the program’s cost test. [Table 25](#) provides an example of the universe of participant and systems costs associated with demand response programs.⁵⁵

⁵⁵ Id at 23.

Table 25: Cost of Demand Response

Type of Cost		Cost	Responsibility/ Recovery Mechanism
Participant costs	Initial costs	Enabling technology investments	Customer pays; incentives may be available from public benefit or utility demand response programs to offset portion of costs
		Establishing response plan or strategy	Customer pays; technical assistance may be available from public benefits or utility demand response programs
	Event-specific costs	Comfort/inconvenience costs	Customer bears “opportunity costs” of foregone electricity use
		Reduced amenity/lost business	
		Rescheduling costs (e.g., overtime pay)	
System costs	Initial costs	Metering/communications system upgrades	Level of costs and cost responsibility vary according to the scope of the upgrade (e.g., large customers vs. mass market), the utility business case for advanced metering system or upgrades, and state legislation/policies
		Utility equipment or software costs, billing system upgrades	Utility typically passes cost through to customers in rates
		Customer education	Ratepayers, public benefits funds
	Ongoing program costs ¹	Program administration/management	Costs are incurred by the administering utility, LSE or ISO/RTO and are recovered from ratepayers
		Marketing/recruitment	
		Payments to participating customers	
		Program evaluation	
		Metering/communication ²	

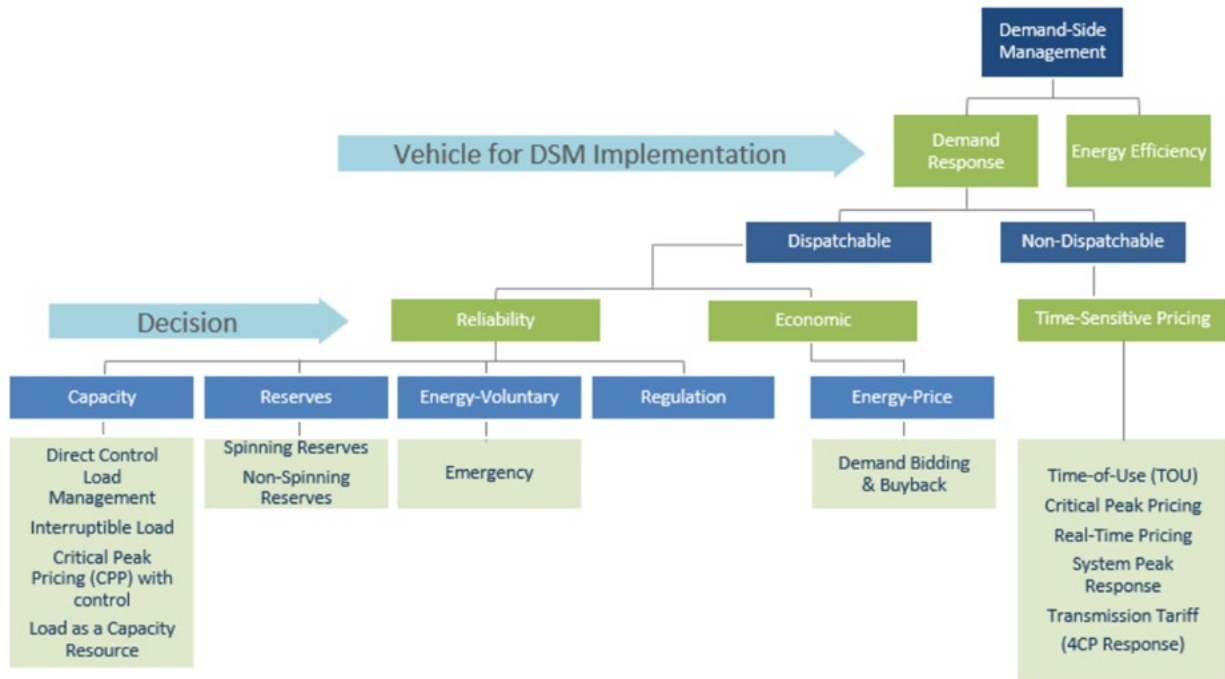
In 2020, FERC issued Order 2222 aimed at promoting “competition in electric markets by removing the barriers preventing distributed energy resources (DERs) from competing on a level playing field in the organized capacity, energy and ancillary services markets run by regional grid operators.”⁵⁶ Existing rules that were developed for traditional resources and commercial and transactional barriers are factors that limited participation of distributed resources and emerging technologies. DERs—including demand response—increase available services in the electricity market and create opportunities for utility programs and program expansion.

Opportunities for demand response resources are expected to grow in the coming years as efforts to decarbonize electricity systems and the wider economy with cleaner, intermittent energy resources are pursued. This anticipated increase in decarbonization and management of intermittent energy resources will also create the opportunity for demand response resources to capture and claim the associated avoided costs while remaining relatively cost neutral. Advances in technology—such as smart controls and access to real-time market data—will also allow distributed resources to act quickly and collectively to address grid needs, thus expanding the quantity and range of services that can be offered. Demand response falls into two general categories: dispatchable and non-dispatchable resources. Figure 28 from NERC and presented in the National Standard Practice Manual for Benefit-Cost Analysis of Distributed Energy Resources highlights classifications, purposes, and example programs of each.⁵⁷

⁵⁶ FERC Order NO. 2222 FACT SHEET. September 17, 2020.

⁵⁷ National Standard Practice Manual for Benefit-Cost Analysis of Distributed Energy Resources. NESP. August 2020. Pg. 7-4.

Figure 28: Types of DR



Source: NERC 2011.

In 2015, The Brattle Group highlighted that in the United States, a “5% reduction in peak demand through DR programs could lead to \$35 billion in savings over a 20 year period.”⁵⁸ In the interim, innovative approaches to utilizing household appliances and to aggregate distributed resources through cloud-based virtual power plant software aggregation systems⁵⁹ have expanded the capabilities in the marketplace, allowing a higher level of demand response in the market for a wider range of services that achieve even greater savings. Aggregation of energy storage systems, cycling of air conditioning systems, and the ability to remotely adjust smart thermostats have increased the flexibility of demand response programs.

Program administrators in Connecticut are incorporating more load flexibility into demand management plans including prioritizing expansion of “active demand response offerings to support electrification and carbon neutrality.”⁶⁰ Electrification and decarbonization efforts have been facilitated in other states, through the energy programs, as system benefits are combined with other value streams to screen program costs and benefits for several years. The synergies are now recognized in program offerings in Connecticut. “For the 2022 – 2024 term, the Companies will promote the co-delivery of energy efficiency and demand management programs that support decarbonization and carbon neutrality, including smart thermostats, electric vehicle chargers, and battery storage.”⁶¹

⁵⁸ Valuing Demand Response: International Best Practices, Case Studies, and Applications. The Brattle Group. January 2015.

⁵⁹ Virtual Power Plant Evaluation: SMUD Board Committee of Energy & Customer Services. EPRI. June 2016.

⁶⁰ 2022-2024 Conservation & Load Management Plan: Connecticut Energy Efficiency & Demand Management Plan.

⁶¹ 2022-2024 Conservation & Load Management Plan: Connecticut Energy Efficiency & Demand Management Plan.

6.5 DEMAND RESPONSE RESOURCES IN OTHER JURISDICTIONS

Demand Response Resources (DRRs) in organized wholesale markets operate under the rules developed by ISOs to ensure system reliability and resource adequacy. ISO market prices send signals in the form of electricity prices to the market; resources can receive compensation for services that balance load. Advances in technology and the deployment of grid-interactive devices like thermostats and inverters allow utilities and third parties to aggregate the reduction or expected reductions in usage from numerous customers to impact the ISOs load forecast.

In California, demand response is incorporated into the state Resource Adequacy (RA) program. The California Public Service Commission (CPUC) establishes the RA obligations for all Load Serving Entities (LSEs) within the CPUC's jurisdiction, including investor-owned utilities (IOUs), energy service providers (ESPs), and community choice aggregators (CCAs).⁶² Resource Adequacy is jointly administered by the CPUC, the CAISO, and other local authorities in the CAISO balancing authority area.⁶³ Resources in the DR programs in California are categorized as dispatchable (supply-side) and non-dispatchable (load-modifying) demand response.⁶⁴ Dispatchable resources participate in the wholesale market or ancillary service market.⁶⁵ On the other hand, load-modifying DR "is typically assigned a value from the reduction in procurement of grid services through the market (e.g. as expressed in capacity procurement needs from long-term load forecasts)."⁶⁶

NYISO offers four demand response programs "to enhance system reliability and reduce overall production costs."⁶⁷ Two of the programs (Emergency Demand Response Program (EDRP) and the Installed Capacity – Special Case Resource (ICAP/SCR)) are designed to reduce power consumption by directing demand response resources to reduce load or to use qualified Local Generators to remove load from the system during grid emergencies or when reserve shortages are anticipated or actually occur."⁶⁸ The other two programs are economic demand response programs— Day-Ahead Demand Response Program (DADRP) in the Energy market, and the Demand-Side Ancillary Services Program (DSASP) in the Ancillary Services market—"flexible loads to effectively increase the amount of supply in the market and moderate Energy prices."⁶⁹

⁶² California Public Service Commission. Resource Adequacy. [Resource Adequacy Homepage \(ca.gov\)](http://www.cpuc.ca.gov/ResourceAdequacy)

⁶³ Supply-Side DR & Stakeholder Working Group Process. CEC Demand Response Team. [TN238933_20210719T114239_Presentation for Supply-Side DR & Stakeholder Working Group Process.pdf](https://www.energy.ca.gov/sites/default/files/2021-07/TN238933_20210719T114239_Presentation%20for%20Supply-Side%20DR%20&%20Stakeholder%20Working%20Group%20Process.pdf)

⁶⁴ [2025 California Demand Response Potential Study - Charting California's Demand Response Future: Final Report on Phase 2 Results \(lbl.gov\)](https://www.lbl.gov/publications/2025-california-demand-response-potential-study-charting-california-s-demand-response-future-final-report-on-phase-2-results)

⁶⁵ 2015 California Demand Response Potential Study – Charting California's Demand Response Future: Interim Report on Phase 1 Results. Lawrence Berkeley National Lab. [2015_dr_potential_study_phase1_final_report.pdf \(lbl.gov\)](https://www.lbl.gov/publications/2015-dr-potential-study-phase1-final-report)

⁶⁶ 2015 California Demand Response Potential Study – Charting California's Demand Response Future: Interim Report on Phase 1 Results. Lawrence Berkeley National Lab. [2015_dr_potential_study_phase1_final_report.pdf \(lbl.gov\)](https://www.lbl.gov/publications/2015-dr-potential-study-phase1-final-report)

⁶⁷ NYISO 2020 Annual Report on Demand Response Programs. [820330e8-d51f-9315-fa01-c6590a62013a \(nyiso.com\)](https://www.nyiso.com/documents/20199/20200/820330e8-d51f-9315-fa01-c6590a62013a)

⁶⁸ NYISO 2020 Annual Report on Demand Response Programs. [820330e8-d51f-9315-fa01-c6590a62013a \(nyiso.com\)](https://www.nyiso.com/documents/20199/20200/820330e8-d51f-9315-fa01-c6590a62013a)

⁶⁹ NYISO 2020 Annual Report on Demand Response Programs. [820330e8-d51f-9315-fa01-c6590a62013a \(nyiso.com\)](https://www.nyiso.com/documents/20199/20200/820330e8-d51f-9315-fa01-c6590a62013a)

Table 26 lists examples of program offerings in other New England states and California, Hawaii, and New York.

Table 26: Overview of Demand Response Programs

New England States	California (SCE) ²	Hawaii (HECO)	New York
Massachusetts <ul style="list-style-type: none"> Residential Wi-Fi Thermostats Residential EV Residential Battery Storage C&I ConnectedSolutions Winter Demand Response New Hampshire <ul style="list-style-type: none"> NHEC Demand Response Program Vermont <ul style="list-style-type: none"> Defeat the Peak Pilot Maine ¹ <ul style="list-style-type: none"> Commercial Battery Storage Load Management Pilot Aggregate Distributed Energy Resource (DER) Load Management Pilot Cold Storage Facility Load Management Pilot Isle au Haut Thermal Energy Storage Load Management Pilot Rhode Island <ul style="list-style-type: none"> C&I ConnectedSolutions 	<ul style="list-style-type: none"> Automated Demand Response (Auto-DR) Control Incentive Critical Peak Pricing (CPP) Real-Time Pricing (RTP) Summer Discount Plan (SDP) Smart Energy Program (SEP) Capacity Bidding Program Time of Use Base Interruptible Program (TOU-BIP) Emergency Load Reduction Program (ELRP) Agricultural and Pumping Interruptible Program (AP-I) 	<ul style="list-style-type: none"> Residential Load Control Water Heater Residential Load Control Air Conditioner Commercial and Industrial Direct Load Control FastDR Pilot Third-Party Aggregation 	<ul style="list-style-type: none"> Commercial System Relief Program Distribution Load Relief Program Auto Dynamic Load Management (Auto-DLM) Term Dynamic Load Management (Term-DLM) Targeted Demand Response (TDR) Program

¹ Efficiency Maine Trust submitted Triennial Plan V for approval – which includes a Demand Management Program.

² PG&E and SDG&E offer a subset of the SCE programs.

Depending on the program, customers can benefit from participating in demand response programs in a number of ways. First, customers can receive direct compensation for reducing system load during demand response events (energy payments). Customers may also receive capacity payments for the nominated capacity declared dispatchable during a curtailment event, even if no dispatch event occurs. Customers will also likely benefit from reduced demand charges if their facility peak is coincident with system peak. Second, and more indirectly, customers also benefit through energy and capacity price suppression stemming from avoided investments in system upgrades provided by available demand response capacity. Wholesale market orientated

programs where capacity is expected to be available day ahead or in real-time (e.g., within 30 minutes) can carry non-performance penalties, where utility driven programs targeting local system peaks or reliability concerns often do not penalize participants for non-performance. Participants contributing smaller capacity nominations to demand response programs will often enroll with a Curtailment Service Provider or Aggregator who will monitor performance, provide dispatch call signals, and process incentive payments on behalf of the participants and in coordination with the utilities and relevant ISO, whereas larger customers can and will self-enroll in the programs of interest.

6.6 AVOIDED ENERGY SUPPLY COMPONENTS (AESC)

UI and Eversource currently utilize the AESC 2021 study's estimate of avoided costs for active demand response programs, including active load management and peak load shifting programs.⁷⁰ The 2021 study included demand response and BTM energy storage measures in active demand management.⁷¹ Customer responsiveness and dispatch approaches impact the value of demand response and should be accounted for in planning, implementation, and evaluation of programs. The value of demand response is limited to the benefits recognized in the cost-effectiveness framework. For the UI and Eversource programs, the value generated by demand response resources largely fall into the following categories:

- **Avoided generation capacity costs:** Program administrators can either bid into the Forward Capacity Auctions (FCAs) (cleared) or reduce peak summer loads through non-bid capacity (uncleared), which then become phased-in load forecasts for subsequent FCAs.⁷² Both UI and Eversource claim uncleared capacity values as depicted in AESC 2021 for all current measures depicted in the provided spreadsheet calculations. "Program administrators can claim avoided capacity by either bidding capacity (cleared) into the FCAs, or by reducing peak summer loads through non-bid capacity (uncleared) (which then becomes phased-in load forecasts for subsequent FCAs)."⁷³
- **Demand Reduction Induced Price Effect (DRIPE):** Effectively suppresses energy and capacity market prices and benefits all ratepayers through lower energy costs.⁷⁴ Both program administrators currently claim Capacity DRIPE uncleared values as depicted by AESC 2021 for all demand response measures. Uncleared DRIPE benefits are claimed by those administrators who do not submit or otherwise do not expect demand response resources to clear in the ISO-NE FCM. "Demand-response and load-management programs that do not clear in the FCM also generate capacity DRIPE benefits, albeit with different timing and of different magnitudes."⁷⁵

⁷⁰ Avoided Energy Supply Components in New England: 2021 Report. Prepared for AESC 2021 Study Group. Released March 15, 2021.

⁷¹ Id at 75.

⁷² Id at 64.

⁷³ Id at 68.

⁷⁴ Id at 66.

⁷⁵ Id at 213.

- Avoided transmission and distribution (T&D) costs for both Pooled Transmission Facilities (PTF) and Local T&D are two additional categories of avoided costs applied to demand response measures by both utilities.
 - Program administrators develop Local T&D values to fully ascribe benefits of demand response, in addition to crediting demand-side measures with value for avoiding T&D costs across a service territory.⁷⁶ Both UI and Eversource have completed specific Local T&D studies and utilized these avoided costs in their benefit-cost analysis tools. Eversource’s most recent Local T&D study was completed in 2018 and for UI in 2017. Both utilities use the same weighted average for these avoided costs across both studies.

As depicted by cost benefit analysis information along with the utility-specific Local T&D values provided, the total value for demand response claimed by the utilities is provided in the table below for the planning year 2022. The amounts listed in the table below and claimed in both utilities separate benefit-cost analysis tools are the same as confirmed by a review of each utility’s provided benefit-cost analysis spreadsheets and reflective of AESC 2021 avoided cost values.

Table 27: Eversource and United Illuminating Avoided Costs for the Planning Year 2022⁷⁷.

Avoided Cost	Avoided Costs (\$/kW)
Capacity, Uncleared	\$60.25
Capacity DRIPE, Uncleared	\$273.48
Reliability, Uncleared	\$28.73
Transmission & Distribution, PTF	\$92.04
Local Transmission	\$0.96
Local Distribution	\$34.52

Eversource’s benefit-cost analysis spreadsheet calculations consider avoided costs related to Electric Energy, DRIPE, and Environmental Compliance when assessing cost-effectiveness for battery storage dispatch measures. However, the accounting for this results in a net penalty (or cost) to the measure meant to represent the additional energy usage resulting from round-trip efficiency losses.

6.7 PROGRAM COST-EFFECTIVENESS METHODOLOGY

The justification for spending on energy conservation programs lie in the objectives of the programs and the evaluation methodology. The program objectives lay the framework for benefits and costs that are acknowledged and pursued and form the basis for cost-effectiveness assessments. Connecticut energy efficiency and conservation programs are evaluated using the

⁷⁶ Id at 258.

⁷⁷ As of December 2021.

Modified Utility Cost Test (MUCT) for non-limited income programs and Total Resource Cost (TRC) test for limited income programs. The MUCT includes utility-specific system benefits, and “avoided costs of oil and propane, and program costs expended to achieve the savings.”⁷⁸ In the 2022-2024 C&LM Plan⁷⁹, the C&I DR cost benefit ratios were above 2.0 for all three years. Residential DR was below 1.2 for all three years.

The TRC, on the other hand, is used to evaluate limited income programs and “includes environmental benefits (such as water savings and reduced air emissions) and maintenance savings.” All costs associated with acquiring the TRC benefits are included in the evaluation including cost to the program administrator and participants.⁸⁰

6.8 ENERGY AND CLIMATE GOALS

Increasingly across the country, states are seeking synergies in attaining energy and climate goals by adapting energy programs to support greenhouse gas reduction. In March 2021, Massachusetts instituted the Climate Act, committing the Commonwealth to achieving net zero emissions by 2030.⁸¹ The Climate Act includes provision for the Secretary of Energy and Environmental Affairs to “set a goal, expressed in tons of carbon dioxide equivalent, every three years for the succeeding Mass Save Energy Efficiency Plans.”⁸² Also in May 2021, the California Public Utilities Commission (CPUC) “reformed its approach to energy-efficiency programs to better align with reducing greenhouse gas (GHG) emissions and support customer equity and long-term energy grid stability.”⁸³

Massachusetts and California energy conservation programs currently recognize energy and non-energy benefits to participants, non-participants, and the electric grid. Both Massachusetts and California evaluate program cost-effectiveness using the TRC as the primary test. The American Council for an Energy Efficient Economy (ACEEE) reported that more than 15 states recognize health and environmental benefits in cost-effectiveness tests.⁸⁴ States that use the TRC and Societal Cost Test (SCT) monetize non-energy benefits such as avoided utility environmental compliance costs (avoided CO₂, NO_x, and particulate matter PM₁₀), societal environmental and/or public health benefits (air and water pollution, greenhouse gas emissions, and cooling water use), and participant health benefits (missed days of work, cold- and heat-related thermal

⁷⁸ Economic Analysis (Cost-Benefit Testing) of Conservation and Load Management Programs: Overview of DEEP’s Framework for Review. CT DEEP Public Information Meeting. September 13, 2018.

⁷⁹ <https://energizect.com/sites/default/files/2022-03/Final%202022%202024%20Plan%20%2803-01-22%29.pdf>

⁸⁰ Economic Analysis (Cost-Benefit Testing) of Conservation and Load Management Programs: Overview of DEEP’s Framework for Review. CT DEEP Public Information Meeting. September 13, 2018.

⁸¹ RE: Greenhouse Gas Emissions Reduction Goal for Mass Save: The Commonwealth of Massachusetts. Executive Office of Energy and Environmental Affairs. July 15, 2021.

⁸² RE: Greenhouse Gas Emissions Reduction Goal for Mass Save: The Commonwealth of Massachusetts. Executive Office of Energy and Environmental Affairs. July 15, 2021.

⁸³ CPUC Better Aligns Energy Efficiency Programs to Reduce GHG Emissions, Support Equity, and Increase Grid Stability. California Public Utilities Commission. May 20, 2021.

⁸⁴ Id at 6.

stress).⁸⁵ A summary of monetized avoided utility environmental and compliance costs used in cost-effectiveness tests in the ACEEE report can be referred to in [Table 28](#).

Table 28: Summary of Avoided Utility Environmental and Compliance Costs⁸⁶

State	Assessment(s) – level	Monetized or proxy variable	Included avoided utility environmental compliance costs
CA	TRC - portfolio	Monetized	CO ₂ , NO _x , and particulate matter (PM ₁₀)
CO	Modified TRC – program	Monetized and proxy	Avoided CO ₂ , SO ₂ , and NO _x emissions.
CT	UCT/TRC income limited programs	Monetized	Eversource uses avoided environmental compliance costs when calculating benefit-cost impacts for summer daily and targeted battery dispatch - these calculations result in a net-negative impact to benefits (i.e., a cost) for these measures. No avoided environmental and compliance costs were claimed in the UI benefit-cost test framework as depicted in the provided benefit-cost spreadsheet.
DE	TRC – program	Monetized	Avoided environmental compliance costs, where such costs can be directly tied to changes in energy use.
ID	UCT/TRC	Proxy	10% conservation benefit adder to calculate cost effectiveness of DSM programs.
IL	TRC-portfolio	Monetized	Recent utility plans include reasonable estimates of financial costs likely to be imposed by future regulations and legislation on emissions of greenhouse gases.
IO	SCT - portfolio, program, project, measure	Proxy	10% externality factor applied to avoided electric capacity costs. Utility may propose a different externality factor but must document its accuracy.
MA	TRC - program	Monetized	Avoided costs used are according to AESC, including cost of complying with the Regional Greenhouse Gas Initiative, NO _x and SO ₂ reduction policies, and state environmental rules.
NH	TRC - program	Monetized	Cost of carbon compliance included in Synapse’s AESC avoided cost values.
NY	SCT + CO ₂ damage costs - portfolio	Monetized	Avoided compliance costs of Regional Greenhouse Gas Initiative and SO ₂ and NO _x cap-and trade markets reflected in locational-based marginal price.
OR	TRC - program	Proxy	Compliance with potential future state carbon policies in addition to 10% adder.

⁸⁵ Id at 6.

⁸⁶ Cost-Effectiveness Tests: Overview of State Approaches to Account for Health and Environmental Benefits of Energy Efficiency. American Council for an Energy Efficient Economy. December 2018.

Connecticut also has ambitious climate goals as the Global Warming Solutions Act required a 80% reduction of greenhouse gas emission by 2050 based on 2001 levels.⁸⁷ The 2018 Connecticut Comprehensive Energy Strategy highlighted several impacts of energy policy choices over the next several years including: reduce peak energy demand, decrease fuel security risk and system costs; reduced cost of renewable energy deployment and integration, and pursue new more sustainable approaches to funding and procuring energy efficiency.⁸⁸ The report included discussion of enhanced opportunities to expand demand response offerings in the summer and winter through ongoing grid modernization efforts.⁸⁹

DR will continue delivering benefits in the current program cycle as lessons learned from the 2019 – 2021 program cycle are implemented to improve active demand response strategies. As stated in the 2022 – 2024 C&LM Plan, the role of DR is expected to “significantly reduce peak demand and greenhouse gas emissions, helping to mitigate the impact that the state’s building sector has on the environment and climate change.”⁹⁰ This could further facilitate the state’s energy and climate goal by recognizing the uncaptured societal benefits achieved through DR programming. All the value streams available to DR necessitate a broader view of benefits of peak reduction from both active and passive demand response and the associated costs to achieve the reductions. States with similarly ambitious climate and energy goals utilize broader cost-effectiveness criteria to access the full suite of value streams available for programs. Although controversial in some cases, these criteria include an expanded view of non-energy benefits. In addition, estimating the value of some non-energy benefits can be complicated, leading many to resist any attempt at monetizing them. Most states that currently account for non-energy benefits typically do so only for benefits that are readily quantifiable.⁹¹ However, these benefits are widely used around the country.

Table 30 highlights some non-energy benefits that could open additional value streams in Connecticut by using a broader cost-effectiveness test for non-limited income programs and provides an overview of the benefits and costs of the PACT/UCT, SCT, and TRC for comparison.⁹² The PACT/UCT allows residents in Connecticut benefit from the aggregation of energy savings and “passive” demand reductions that are bid into ISO-NE’s FCM.⁹³ In states that use the SCT and TRC, market revenues play an important role in creating additional value streams for the residents through financial compensation and market mechanisms to monetize benefits beyond those to electric system. However, the final decision on whether to accept, and how to incorporate, NEIs is determined by state policy. Aligning these additional program benefits where possible and aligning with jurisdictional policy goals can also foster synergies in program

⁸⁷ Comprehensive Energy Strategy. CT General Statutes Section 16a-3d. February 8, 2018.

⁸⁸ Id at 13.

⁸⁹ Id at 63.

⁹⁰ 2022-2024 Conservation & Load Management Plan: Connecticut Energy Efficiency & Demand Management Plan.

⁹¹ Cost-Effectiveness Tests: Overview of State Approaches to Account for Health and Environmental Benefits of Energy Efficiency. American Council for an Energy Efficient Economy. December 2018.

⁹² Benefit-Cost Analysis; Docket No. 17-12-03(RE03). Guidehouse. October 1, 2021

⁹³ 2022-2024 Conservation & Load Management Plan: Connecticut Energy Efficiency & Demand Management Plan.

value creation and policy impact. Considerations of incorporating hard to quantify metrics as listed by NEEP and utilized in other states include:

- Assign a monetized value for use when an exact cost can be identified. A monetized value can come from a variety of sources such as studies of the costs of arrearages to consumers and utilities, third-party models and engineering, surveys, and other state or national regulations.⁹⁴
- Assign an adder for difficult or costly to monetize benefits. This can also be used to combine multiple impacts into one factor. Adders can also be applied across a range of benefits. Currently, five percent, ten percent, and 15 percent adders have been used to account for low-income benefits, emissions reductions, and economic benefits in various states.⁹⁵

Table 29: Benefits and Cost of PACT/UCT, SCT, and TRC

Category	PACT/UCT	SCT	TRC
Avoided Energy*	Benefit	Benefit	Benefit
Avoided Generation Capacity*	Benefit	Benefit	Benefit
Avoided T&D Capacity*	Benefit	Benefit	Benefit
Reliability*	Benefit	Benefit	Benefit
DRIFE Energy Impacts*	Benefit	Benefit	Benefit
DRIFE Capacity Impacts*	Benefit	Benefit	Benefit
Cross-DRIFE Impacts*	Benefit	Benefit	Benefit
Non-Embedded Emissions		Benefit	
Market Revenue		Benefit	Benefit
Avoided Ancillary Services		Benefit	Benefit
Job Creation Benefits		Benefit	
Net Societal Non-Energy Benefits		Benefit	Benefit ⁹⁶
Upfront Program Incentives*	Cost		
Performance Incentives*	Cost		
Upfront Incentive Administration*	Cost	Cost	Cost
Performance Incentive Administration*	Cost	Cost	Cost
Participant Incremental DER Costs		Cost	Cost

* Benefits and Cost accounted for in Connecticut program evaluation

⁹⁴ Implementation Guide: Establishing a Jurisdiction-Specific Cost-Benefit Test. Northeast Energy Efficiency Partnerships. June 2021.

⁹⁵ Id.

⁹⁶ Benefits include: Non-Energy Impacts, Non-Embedded Emissions, and Fossil Emissions. 2022-2024 Connecticut Conservation & Load Management Plan: Connecticut Energy Efficiency & Demand Management Plan.

6.9 CONCLUSION

The value of demand response lies in cost-effectively managing supply and demand on the grid through direct control of end-use or through the influence of market signals. The operationalization of demand-side resources to increase or decrease load through grid services allows quantifiable changes to be compensated in the energy, capacity, and ancillary service markets. Historically, the value of demand response has been recognized in grid services decreasing load. As markets have evolved and technologies have advanced, demand response resources can also absorb excess supply or be dispatched to balance load.

Demand response programs in Connecticut have provided benefits to residents in the state over the past two decades. The rising cost of energy and technological advancements have created opportunities to pursue untapped value streams to manage increased electricity demand and service costs. The pilots conducted by the UI and Eversource provide insights into design and requirements for successful implementation of full-scale demand response programs. Demand response also plays a critical role in reaching the state's ambitious climate and energy goals. For example, one attribute is avoided generation due to reduced system demand. The 2022 – 2024 CL&M plan prioritizes equity, decarbonization, energy affordability, and the role the demand response plays in those efforts. In the long run, there will be opportunities to align the cost-effectiveness screening with the policy objectives through broadening of the recognized benefits. In the short run, increase participation in the active demand response programs would increase the benefits associated with demand reduction; namely, avoided system costs and environmental impacts. Benefits like greenhouse gas emissions reduction and cost of environmental compliance can lower the cost to the utilities and ratepayers and support state climate and energy goals. All residents in Connecticut benefit from this reduction, and these benefits provide value that can be captured in program evaluation.

Demand response resources also provide benefits that are not easily quantifiable and do not directly improve system reliability or reduce system costs. Non-energy benefits can create value that are not currently compensated through market mechanisms. Non-energy benefits from DR are typically not observed due to the short duration of events, though it can support renewable energy when present on the generation stack margin. Hence, non-energy benefits such as reduction of negative environmental impacts and improved health outcomes are recognized in some cost-effectiveness evaluations and provide long-term benefits to residents. For example, avoided generation abates environmental pollution for residents in communities surrounding the power plant, improving their long-term health outcomes and abating long-term impacts of carbon dioxide on the environment. Over 15 states across the country recognize and monetize non-energy benefits—such as health and environmental—to facilitate climate and energy goals through expanded program options. States such as Massachusetts and California evaluate cost-effectiveness of energy conservation programs using the TRC. Adoption of the TRC for the non-limited income program would not present significant administrative barriers as program administrators, evaluators, and regulators are familiar with the use of the TRC test for existing limited income programs.

6.10 RECOMMENDATIONS FOR FUTURE PROGRAM DESIGN

The planning and implementation of future programs could benefit from the following:

- **Recommendation 1: We recommend the state encourage the possibility of utilizing a different cost-effective test, which is a state level decision.** Monetizing and incorporating additional value streams including avoided environmental and compliance costs as seen in other states can boost program cost-effectiveness and help meet state policy goals like decarbonization and enhanced environmental quality. Regulators direct utilities in other states to monetize or use proxy values to account for a variety of environmental externalities. Additional avoided costs that reflect reduced environmental impact or increased human wellbeing in the primary test can reveal potentially new value streams but require a benefit cost framework allowing them.
- **Recommendation 2: Consider bidding DR resources into the ISO-NE marketplace.** By not participating in the ISO-NE market, UI and Eversource forego a base payment for availability and a pay-for-performance payment when ISO-NE calls on resources to reduce load. Instead, program economics rely on the premise that peak shaving will lower their peak load forecast and future capacity obligations. If the EDCs opt to pursue wholesale recognition of their DR programs, there are ways to mitigate participation risks (e.g., using qualified Curtailment Service Providers or aggregators) to make the risk profile of participating more acceptable.

Section 7 Study Implications for the Connecticut PSD

Key Findings

- Demand response measurement and reporting guidelines are not clearly defined in Connecticut
- The PSD is a potential venue to memorialize some foundational elements even if no measures are added

While the value of demand response evaluations as ad hoc processes to determine program impacts is clear, there is more value in defining standard processes to regularly estimate program capability and incorporate those impacts into utility planning. In Connecticut, the Program Savings Document (PSD) provides guidance on how to calculate savings for energy efficiency programs but does not discuss processes for assessing demand response savings at all.

As part of the X1932 Study, the evaluation team investigated how different jurisdictions included demand response evaluation methods, impacts, and reporting frequencies in state reporting and planning documentation. The goal of this task was to determine:

- What do other jurisdictions include?
- What types of DR guidance could utilities, regulators, and other stakeholders use if DR measures, or program archetypes was included in the PSD?
- Should there be DR-specific guidance included in the Connecticut PSD?

The following sections discuss the findings of the evaluation team and provide recommendations for materials that may be of use in planning for demand response initiatives as they scale and evolve.

7.1 COMPONENTS FOR CONSIDERATION IN PSD GUIDANCE

Different jurisdictions include different components in their treatment of demand response impacts. While the frequency and level of detail required in different jurisdictions vary, methods and values typically fall in to two main categories:

1. **Ex Post Reporting:** requirements that define methods and reporting for events that occurred over the program's operating season. This typically includes quantifying impacts for each hour of each event day and summarizing overall performance of historical events⁹⁷. Guidance for ex post impact estimation could include deemed savings, meter-

⁹⁷ For example, see the treatment of ex post impacts for each event day in the following series of tables http://calmac.org/publications/SCE0451.02_PY2020_SCE_SEP_Ex_Post_Load_Impacts_Public.xlsx

based methods such as day- or weather-matching customer baseline using a control group as part of a randomized control trial or a quasi-experimental matched control group, or a regression-based approach using participant data only.

2. **Ex Ante Reporting:** requirements that define production of estimates of what load impacts a program could provide under standard planning conditions. To produce ex ante estimates, it is necessary to define these planning conditions up front, and to determine what data, if any, from an ex post assessment is included in the ex ante estimates. For example, California produces ex ante estimates for each demand response program in the state from 4 PM to 9 PM under weather-normalized peak conditions defined using 1-in-2 and 1-in-10 weather years⁹⁸. This can be accompanied or substituted by a time-temperature matrix that is useful for utility operations planning beyond long term capacity planning that the 4 PM to 9 PM average is used for. Simpler approaches can be used, like a simple scalar weather normalization procedure. Vertically integrated utilities that complete Integrated Resource Plans every two to four years typically estimate DR capability under conditions that mirror their peak load forecast.

7.1.1 Performance Assessment Process

The PSD may also offer guidance in assessing demand response performance. Depending on the region, the evaluation team found that different groups may produce estimates of ex post and ex ante impacts. In some jurisdictions, program performance may be based on estimates of ex post estimates. Ex post impact assessments may be done by the ISO or by the implementing vendor, an EM&V contractor, or simply by the utility applying algorithms defined in the TRM or PSD.

Alternatively, some jurisdictions define capacity and performance based on ex ante estimates. In these cases, there is typically a process by which ex post results are incorporated into the ex ante estimates. Transforming these ex post impacts to a value for claimed savings can vary. Annual reporting of program impacts can provide a basis for claimed savings, but ultimately the definition of claiming conditions matters as much as the specifics of how ex post impacts are translated to an ex ante value.

Claiming conditions vary by state. Options include:

- Average performance on monthly ISO-NE peaks?
- Weather-normalized performance for an event under certain daily maximum temperature conditions or at the system peak hour(s).
- Average ex post performance for events that meet a certain temperature threshold

⁹⁸ An example of ex ante load impacts for a DR program can be found at the following link <http://calmac.org/publications/SCE0451.03 PY2020 SCE SEP Ex Ante Load Impacts Public.xlsx>

The appropriate claiming conditions for Connecticut are difficult to assess because the theory whereby DR programs generate capacity and other benefits is not clearly documented in the C&LM Plan. Because the DR programs are not recognized at the wholesale level, the underlying premise is that peak shaving will place downward pressure on the peak load forecast and lower capacity obligations for the EDCs and their ratepayers. Given this theoretical basis for operating DR programs, the C&LM Plan should include a detailed discussion of the peak load forecasting process and how the EDCs can best place downward pressure of the forecast through DR dispatch.

7.2 REVIEW OF OTHER JURISDICTIONS

As there appears to be no standard DR evaluation framework in Connecticut, the evaluation team investigated the requirements across seven other states: California, Pennsylvania, Massachusetts, Texas, Illinois, New York, and New Mexico. These states all have robust energy efficiency and demand response program accounting rules, including some with market-integrated demand response. Other states were chosen for geographic proximity to Connecticut or due to their participation in ISO-NE. In selecting these states, the evaluation team takes no position on whether any approach should be reapplied directly to Connecticut and intend to illustrate a range of possible approaches with the examples below.

In general, the evaluation team found that guidance for demand response impact estimation came from one or more of three locations:

1. **Technical Reference Manual:** algorithms to determine program impacts could be defined in the state's TRM or its equivalent. These entries could be as simple as defining deemed DR savings or could include more detailed calculation steps. In some cases, the TRM simply refers to savings algorithms that are defined in other locations such as in an evaluation framework or protocol, or through ISO-defined approaches.
2. **PUC-defined Regulatory Protocols:** in some states, evaluation activities are more precisely defined, either through load impact protocols or frameworks that require demand response programs to be evaluated regularly using standard reporting methods and tools. In some cases, these protocols require impacts to be produced for both ex post scenarios (an assessment of what occurred on each event day over the reporting period) and ex ante scenarios (an assessment of program capability under a standard set of planning conditions).
3. **Independent System Operator:** The ISO may provide algorithms, such as baseline methods, to determine the amount of delivered load relief and compensate market participants accordingly. Only programs that are recognized as wholesale resources in the market were subject to ISO-defined methods for estimating program capability.

Table 31 presents a summary of each state's approach.

Table 30: Summary of Methods Used to Assess DR Impacts by State

State	Guidance Provided By			Requirements	
	TRM	PUC	ISO Guidance	Reporting Type	Claimable Reporting Conditions
CA		Load Impact Protocols	Settlement is done using CBLs independent of Load Impact calculations Not all programs are bid into CAISO	Impact Evaluation (Every program, every year, by April 1 st)	Predict 4-9pm aggregate portfolio impacts under utility-specific and CAISO specific peaking conditions (1-in-2 weather year, 1-in-10 weather year)
PA	Basic information in TRM, cite evaluation framework for program-specific guidance	Evaluation Framework		Impact Evaluation (Every program, every year, by January 15 th)	Average Performance across all event hours in program year. Events are triggered by Day Ahead Forecast \geq 96% of year's peak load forecast
MA	Vendor-provided estimates, calibrated by impact evaluation			Impact Evaluation	
TX	Impacts estimated by day matching baselines			TRM-Defined Baselines	
IL	Methods separate for ComEd (in PJM market) compared to other EDCs that are in MISO.		PJM measures compliance against commitments	PJM-Defined CBL	
NY			DR programs that are bid into the market have compensation mechanisms based on performance	NYISO-Defined CBL	
NM		Annual EM&V plan		Impact evaluation (Every program, every year by March 15 th)	Evaluation outputs are IRP inputs. Load reduction at planning conditions.

7.3 RECOMMENDATIONS

The evaluation team recommends that standard reporting methods and criteria be defined in Connecticut for quantifying the value of demand response programs. Nevertheless, the evaluation team does not put forward a strong position on which approach described above should be adopted. In general, some trends emerge from this review of other jurisdictions. First, only programs that are bid into the RTO/ISO market are subject to estimation methodologies defined by ISO rules. Second, most states that have PUC-defined load impact protocols or evaluation frameworks produce annual impact evaluations. Third, methods or estimates defined by the TRM tend to be simpler than other approaches, where deemed savings or a scalar adjustment to deemed savings is used.

With those findings in mind, the evaluation team believes it would be useful to highlight several areas where clarity would be useful for future DR guidance:

- How is peak demand reduction capability defined for active DR resources that are not nominated into ISO-NE? Simply put, what does a MW DR resource mean in terms of season, weather conditions, and time of day? Is a MW at Eversource the same as a MW at United Illuminating?
 - Programs that bid into ISO-NE are already subject to customer baselines and performance mechanisms.
- What is the default connected load per device for mass market programs like connected thermostats, HPWH, and Window AC units?
- What is the default baseline? The evaluation team found that the ISO-NE adjusted 10-of-10 it worked well for a range of demand response programs in Connecticut.
- Are there weather or other planning conditions that define the peak and peak reduction capability? Is there benefit to relying on similar conditions to ISO-NE or would something like a time-temperature matrix be more useful to the Connecticut utilities?

Appendix A UI Smart Savers Baseline Accuracy

Developing an unbiased prediction of what load would have been absent a demand response event is essential to producing a defensible demand response impact estimate. If the methodology tends to produce unbiased estimates of load (i.e., average error of 0), then demand response impact estimates will also be unbiased. If the baseline tends to over predict or under predict load, then demand response impacts will be overstated or understated.

Assessing the accuracy of a baseline on an event day is not possible because the counterfactual is unknown. In other words, we do not know what the demand would have been if the event was not called. However, on non-event, non-holiday weekdays there is no demand response, so using the same algorithm to generate a baseline should reasonably predict the metered load. For these days, the true impact of demand response is 0 kW (or minutes when using runtime), so if the baseline yields a non-zero impact estimate, it can be attributed to error. Individual errors are expected as the lookback window is not intended to be a perfect predictor of future load. That said, an unbiased baseline methodology should produce a distribution of errors which are centered around zero, on average.

We used this knowledge of central tendency of the error to assess the accuracy of the ISO-NE baseline used to construct impacts. More generally, this assessment is performed to understand how the ISO-NE baseline performs for this program, and whether we feel confident in implementing this baseline for our impact evaluation. Since the UI program is not recognized as an active DR resource by ISO-NE, the use of this specific baseline methodology is more of a regional convention rather than a required settlement baseline.

The concept of accuracy can be split into two key metrics: bias and fit. [Table 32](#) summarizes the key metrics for bias and fit that were produced throughout this evaluation.

Table 31: Bias and Fit Metrics

Metric	Description	Mathematical Expression
% Bias	This metric indicates the percentage by which the measurement, on average, tends to over or underestimate the true value. A negative value indicates a tendency to under-predict and a positive value indicates a tendency to over-predict.	$MPE = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)}{\bar{y}}$
CVRMSE	The root mean squared error is a measure of how far the estimated values are from actual values. Technically, it measures the spread of errors, weighing bigger errors more heavily than small ones. The RMSE is normalized by dividing it by the average of the actual values.	$CVRMSE = \frac{\sqrt{\frac{\sum (\hat{y}_i - y_i)^2}{n}}}{\bar{y}}$

Across each year, we selected five non-event weekdays, based on the maximum temperature, to create a set of placebo event days. These days are intended to mimic the demand and load shape of event days, but crucially, are not perturbed like event days. Using the maximum temperature for each eligible placebo days, the five hottest days within each year were selected and are displayed in [Table 33](#).

Table 32: Placebo Event Day Characteristics

Placebo Date	Max Daily Temperature (F)	AMI	Runtime
6/27/2019	87	✓	
6/28/2019*	85	✓	
7/16/2019	87	✓	
7/29/2019	90	✓	
7/31/2019	86	✓	
7/27/2020*	91	✓	✓
7/28/2020	89	✓	✓
7/30/2020	88	✓	✓
8/10/2020	88	✓	✓
8/25/2020	89	✓	✓
6/28/2021	88		✓
6/30/2021	91		✓
7/7/2021	88		✓
7/16/2021*	88		✓
8/27/2021	87		✓

* ISO-NE System Load Peak Day

[Figure 29](#) shows the five placebo events days plotted over the average event day from the AMI data and the runtime data. The average event day loads outside of the event hours are comparable to the various placebo event days. For the 2020 DR season, this event day average includes the two events where all thermostat manufactures were dispatched and the two events where OEM C devices were excluded.

Figure 29: Placebo Event Day Loads



To estimate the accuracy of this baseline method, we implemented the ISO-NE baseline on these placebo event days and investigate the observed load, the baseline, and the error. Since no demand response occurred, the impact estimate (difference between the baseline and the observed load) should be zero and is thus labeled as error. These results for the ISO-NE baseline, aggregated by year, are shown in [Table 34](#). To compare the error across both the AMI and runtime, we use the percent bias. This comparison makes evident that the accuracy of the ISO-NE baseline differs across both years and methods, with the largest differential coming from AMI in the 2019 season. Across the 2020 season, the baseline appears to perform well when compared to the other years. This is important to note because we use the AMI and runtime from 2020 to construct our connected load assumption.

Table 33: Baseline Accuracy Results

Year	Hour	AMI (kW)			Runtime (minutes)		
		Observed Load	Baseline	Bias (%)	Observed Load	Baseline	Bias (%)
2019	16	2.53	2.22	-12.9%	-	-	-
	17	2.77	2.47	-10.7%	-	-	-
	18	2.91	2.67	-7.9%	-	-	-
	19	2.92	2.80	-3.0%	-	-	-
Average	-	2.78	2.54	-8.7%	-	-	-
2020	16	3.04	2.93	-3.7%	29.60	29.26	-0.9%
	17	3.31	3.17	-4.0%	30.59	31.01	2.1%
	18	3.43	3.52	3.5%	30.39	31.98	6.5%
	19	3.43	3.55	4.4%	29.47	31.56	8.7%
Average	-	3.30	3.29	0.1%	30.01	30.95	4.1%
2021	16	-	-	-	27.58	25.52	-7.7%
	17	-	-	-	29.30	26.91	-8.4%
	18	-	-	-	30.26	27.58	-9.0%
	19	-	-	-	29.01	27.59	-4.7%
Average	-	-	-	29.04	26.90	-7.4%	

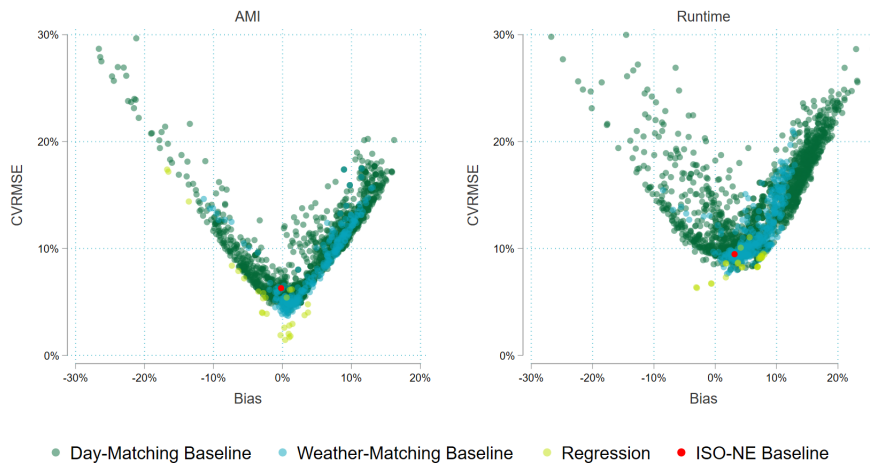
The difficulty with selecting a baseline method is two-fold. Since we have runtime data for 2020 and 2021 but only have AMI data for 2019 and 2020, it is unlikely that the same model will perform the best across the two data types and three years. To understand how this model performs in a broader context, we tested 2,172 additional baseline methodologies and plotted the percent bias and the CVRMSE in a scatterplot. The best models will have a percent bias and a CVRMSE close to zero. Figure 30 plots these metrics during the 2020 DR season for day-matching⁹⁹, weather-matching¹⁰⁰, and regression baselines¹⁰¹.

⁹⁹ This approach relies on electricity use in the days leading up to the event to establish the baseline. A subset of non-event days close to the event day is identified. The electricity use in each hour of the identified days is averaged to produce a baseline.

¹⁰⁰ The process for weather matching baselines is similar to day-matching except that the baseline load profile is selected from non-event days with similar temperature conditions.

¹⁰¹ Regression models quantify how different observable factors such as weather, hour of day, day of week, and location influence energy use patterns. They can replicate many of the elements of day and weather matching baselines.

Figure 30: Bias and Fit Metrics Across Data Types in the 2020 DR Season



This figure seems to indicate that, in both the AMI and runtime data, the ISO-NE baseline is a contender based upon percent bias but is not the most precise method based on CVRMSE. Additionally, the regression-based baselines seem to perform well across both metrics. To understand how the ISO-NE baseline performs against the best regression baseline, we plot both on a single placebo event day in the 2020 season.

Figure 31: ISO-NE Baseline and Regression Baseline on 2020 Placebo Event Day

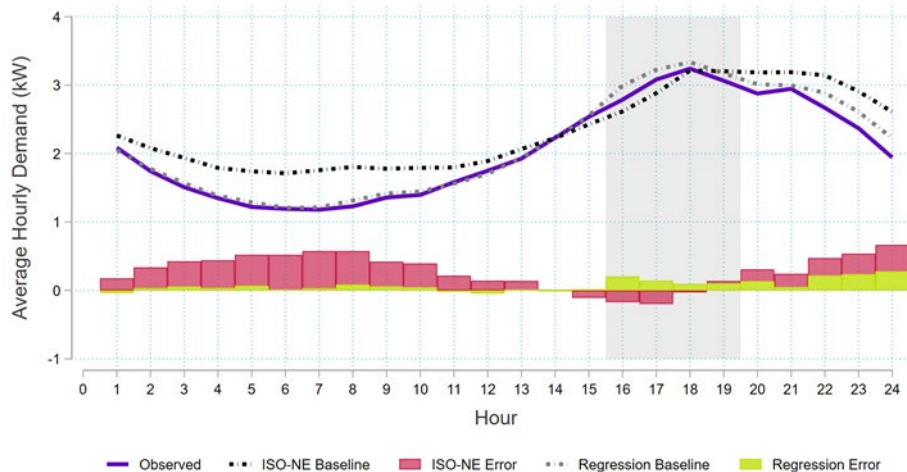


Figure 31 indicates, that during event hours, the baselines tend to perform similarly, with the ISO-NE baseline slightly under-predicting and the regression baseline slightly over-predicting. Balance must be maintained between accuracy and transparency. For example, to continue with the next-day reporting EnergyHub provides, there are limited baselines. This limitation is due to two factors:

1. **Day availability.** The regression-based approach uses data from both the past and the future, not all of which is available for next-day reporting.
2. **Mathematical computations.** Day averaging is computationally less complex than regression analysis. While EnergyHub may have the capability to store regression

coefficients and predict from them, the operational challenges rules for developing and updating the regression coefficients would need to be worked out.

With this in mind, we believe the ISO-NE baseline performs reasonably well and should continue to be implemented given the simplicity and regional acceptance.

Appendix B UI C&I Auto DR Regression Specifications

This appendix provides the details of the regression specifications tested.

Basic Model

The most basic model modeling approach for demand response programs limits model data to non-event and non-weekend/holiday days. This approach removes weekends and holidays that are not accessible to the program and does not directly estimate event reductions as part of the regression model. Such a model specification for this set of hourly load data includes an hourly weekday baseload shape and an hourly weather-correlated effect. The following model provides this specification for a single season of load date (e.g., Summer) where T_{id} represent some form of weather data.

$$L_{idh} = \alpha_{ih} + \beta_{ih}T_{id} + \varepsilon_{idh}$$

Where:

- L_{idh} = Load for customer I for day d and hour h
- T_{id} = Temperature for day d , usually daily average, degree-days, or multi-hour lagged moving average
- α_{ih} = Intercept in hourly baseload shape
- β_{ih} = hourly temperature trends
- ε_{idh} = error

The basic model assumes that a single baseload shape can represent individual weekday baseload shapes across the full timespan of the data. Similarly, weather-correlation effects are linear and do not vary across day of week or across the timespan of the data. These assumptions can be problematic for a particular customer. Within limitations driven by degrees of freedom, we can create models that are flexible to these issues.

Weekday Baseload Model

The weekday baseload model allows a unique baseload shape for each day of the week. For a customer with different schedules of non-weather-correlated production, for example, this model will provide a better weekday specific baseline. The single set of hourly weather-correlated effects remains fixed across weekdays.

$$L_{idh} = \alpha_{idh}D_d + \beta_{ih}T_{id} + \varepsilon_{idh}$$

Where:

- D_d = Weekday dummy
- α_{idh} = Intercept in weekday-specific, hourly baseload shapes

Monthly Baseload Model

The monthly baseload model allows the 24-hour baseload shape to vary month to month through the summer but not across weekdays. It is common for customer load dynamics to vary across

summer months. This specification is flexible to that likely weather-oriented variation and captured in the baseload.

$$L_{idh} = \alpha_{imh}M_m + \beta_{ih}T_{id} + \varepsilon_{idh}$$

Where:

- M_m = Monthly dummy
- α_{imh} = Month-specific, hourly baseload shapes

Basic Model with Squared Weather

Adding an hourly, squared temperature term is a low degrees-of-freedom way of giving the weather trend non-linear flexibility. This specification adds the squared weather term to the basic model.

$$L_{idh} = \alpha_{ih} + \beta_{ih}T_{id} + \beta_{ih}^2T_{id}^2 + \varepsilon_{idh}$$

Where:

- T_{id}^2 = Temperature variable for day i, *squared*
- β_{ih}^2 = hourly squared temperature effect

Weekday Model with Squared Weather

This specification adds the squared weather term to the weekday model.

$$L_{idh} = \alpha_{idah}D_d + \beta_{ih}T_{id} + \beta_{ih}^2T_{id}^2 + \varepsilon_{idh}$$

Weekday Model with Squared Weather and Monthly Dummy

The weekday model with squared weather and monthly dummy adds a monthly dummy to the specification. Unlike the monthly baseload model, this is not a fully monthly 24-hour load shape but a fixed monthly shift that provides flexibility while using minimal degrees of freedom.

$$L_{idh} = \alpha_{idah}D_d + \gamma_{im}M_m + \beta_{ih}T_{id} + \beta_{ih}^2T_{id}^2 + \varepsilon_{idh}$$