|  |  |
| --- | --- |
| To: | Kiersten Williams, EversourceRobert Wirtshafter, EEAC & Lisa Skumatz, SERA |
|  |  |
| From: | Ethan Young, Jackson Lines, & David Basak, GuidehousePace Goodman, ILLUME |
|  |  |
| Date: | May 11, 2023 |
|  |  |
| Re: | CT Study R2212: Behavioral Program Impact Evaluation (Multi-State)PY2022 Delivered Energy Insights (DEI) Program Impact Results  |

# Executive Summary

Eversource administers a behavior program, Delivered Energy Insights (DEI), in Connecticut, Massachusetts, and New Hampshire. The DEI program was designed to encourage program participants to be more energy efficient by sending digital letters via email with personalized energy usage information and energy-saving tips. As such, the program is intended to achieve direct savings through behavioral change.

In this memorandum, Guidehouse and ILLUME (the evaluation team) provide the second-year evaluation results for this program. The evaluation results provide definitive electric impact estimates for New Hampshire. However, for Connecticut and Massachusetts, the evaluation results are not definitive, and the evaluation team encourages stakeholders in each state to determine the appropriate course of action for claiming savings. For example, stakeholders could decide to (a) accept the results as the best estimate of savings, or (b) determine that the uncertainty is too high and savings should be considered zero. Similarly, given the uncertainty in the results, stakeholders may decide to exclude the uplift adjustment for any savings claimed.

In the sections below, the evaluation team provides an introduction, a description of the research approach, research results, and key findings and recommendations.

# 1. Introduction

Since 2021, Eversource has been piloting a new behavior program, Delivered Energy Insights (DEI), in Connecticut, Massachusetts, and New Hampshire. This program was launched in 2021, following the 2019 discontinuation of the Home Energy Reports (HER) program, a residential behavioral program that was implemented by Oracle. The DEI program was designed to encourage program participants to be more energy efficient by sending digital letters via email with personalized energy usage information and energy-saving tips. As such, the program is intended to achieve direct savings through behavioral change.

Eversource emailed energy insight reports on a regular basis to a randomized treatment group in Massachusetts, Connecticut, and New Hampshire throughout 2021 and 2022. The evaluation team previously provided energy savings results from program year 2021 (PY2021) in Connecticut and Massachusetts. For PY2021, the evaluation team did not find statistically significant savings estimates for both Massachusetts and Connecticut. Eversource is now interested in understanding energy savings from PY2022 of its DEI program for Connecticut, Massachusetts, and New Hampshire. With this context, the PY2022 results provide new insight into the DEI savings in New Hampshire and determine whether the DEI savings in Massachusetts and Connecticut increased in the second year. In some cases, behavioral programs achieve greater savings in the second year along an arc as savings ramp up before stabilizing and ultimately declining over time.

#  2. Research Description

The evaluation approach includes two basic components: the study design and uplift adjustment.

***2.1 Study Design***

Eversource administers the program as a randomized control trial (RCT), where study participants are randomly assigned to treatment and control groups. In this study design, Eversource sends personized DEI messages to the treatment group and does not send any DEI messaging to the control group. As such, evaluators use pre- and post-intervention usage data within regression modeling to determine savings, where the control group provides an ideal baseline as a component of the difference-in-difference modeling approach. Evaluators and researchers typically consider RCT a gold-standard for study design because it provides unbiased, accurate results.

In subsequent sections, the evaluation team provides PY2022 evaluated savings estimates by state, fuel type, and whether participants were a part of the discontinued HER program. The evaluation team also provides exploratory savings estimates separately for DEI participants in the lower, middle, and upper 33% of energy users.

In the appendix, the evaluation team describes the methodology in greater detail, which includes results and descriptions for (a) verifying equivalency between the treatment and control groups, and (b) the difference-in-difference modeling approach.

***2.2 Uplift Adjustment***

Behavioral programs encourage energy efficient behaviors and choices via messaging. In some cases, those behaviors and choices can include savings that are claimed in other programs (e.g., upgrading windows or taking advantage of weatherization services). Evaluators typically use uplift adjustments to correct this issue and avoid double counting savings. More specifically, the uplift adjustment removes any savings from the behavioral evaluation that are claimed through other programs.

For PY2022, the evaluation team and stakeholders agreed to use a deemed uplift adjustment of 0.02% of baseline usage based on PY2021 results and recommendations. However, the evaluation team provides savings results prior to the uplift adjustment should stakeholders decide to exclude the uplift adjustment for any savings claimed.

In the PY2021 report, the evaluation team recommended using a deemed uplift adjustment for several reasons. First, the evaluation team did not find statistically significant DEI savings for Massachusetts and Connecticut in PY2021 prior to the uplift adjustment. As such, the evaluation team prioritized resources towards estimating savings prior to the uplift adjustment in PY2022. Second, the uplift results from PY2021 were highly uncertain yet were, as a percentage of baseline usage, consistent with past HER evaluations and generally a small percentage of the expected savings from behavioral programs. Given this finding, the evaluation team did not expect to be able to provide more accurate uplift results that would be substantially different for PY2022, where any differences could likely reflect the high degree of uncertainty rather than true changes in uplift over time. Lastly, using deemed uplift values still provides a mechanism to avoid double counting savings. In summary, the evaluation team recommended using deemed uplift values for PY2022 to shift evaluation resources toward the most impactful research topics, while still avoiding double counted savings.

Once determining that a deemed uplift value could be appropriate for the PY2022 evaluation, the evaluation team calculated an uplift of 0.02% of baseline usage, which reflects an average across gas and electric uplift results in Massachusetts and Connecticut. The evaluation team was comfortable recommending this value for PY2022, because it reflected an average within the range of relevant studies and would be a small percentage of typical behavioral program savings, which can range from 0.15% to 0.75% of baseline usage.

As a cautionary note, uplift effects can change as behavioral programs mature and the evaluation team encourages Eversource to revisit uplift research periodically to determine if updates to the uplift adjustment are warranted.

# 3. Research Results

In following subsections, the evaluation team provides the DEI PY2022 evaluated savings estimates by state, fuel type, and whether participants were a part of the discontinued HER program.[[1]](#footnote-2) The evaluation team also provides exploratory savings estimates separately for DEI participants in the lower, middle, and upper 33% of energy users. We present results as percent savings, total savings[[2]](#footnote-3) prior to uplift and the associated confidence intervals, number of treatment accounts, uplift, and savings after the uplift adjustment.

It’s important to note that nearly all results are not statistically different than zero. Because the study is designed as a randomized control trial (RCT), the regression results can be considered unbiased and that they provide our best estimates of impacts and other results. However, given the uncertainty in the results, in most cases we cannot say with 90% confidence that the true impacts are above or below zero. The evaluation team provides all results for due diligence and documentation, and also to enable stakeholders in each state to determine the appropriate course of action for claiming savings (e.g., to accept the results as the best estimate of savings or to determine that the uncertainty is too high and savings should be considered zero). The following discussion of results notes this as well, but it also includes discussion of point estimates. The evaluation team urges caution and transparency if using the point estimates for reporting purposes.

***3.1 Massachusetts Electric Savings***

Table 1 illustrates estimated electric savings attributed to the DEI program in Massachusetts. PY2022 electric savings in Massachusetts were approximately 0.01%, amounting to 108,646 kWh across all DEI participants; however, results were not statistically significant. When including uplift, overall savings were -202,800 kWh.

Table . PY2022 Massachusetts DEI Electric Savings

| Savings Category | Verified Net Savings, Prior to Uplift Adjustment |  | Savings Uplift in Other EE Programs, kWh | DEI Savings net of Uplift, kWh |
| --- | --- | --- | --- | --- |
| % | kWh | 90% Confidence Interval (kWh) | Treatment Count |
| Overall | 0.01% | 108,646 | -2,814,785/3,032,076 | 175,958 | 311,445 | **-202,800** |
| **Prior HER Status** |  |  |
| HER Recipient | -0.17% | -1,363,516 | -3,340,151/613,118 | 88,746 | 163,222 | **-1,526,738** |
| HER Control | 0.15% | 387,023 | -757,054/1,531,101 | 26,679 | 51,817 | **335,207** |
| No HER engagement | 0.31% | 1,440,944 | -365,986/3,247,874 | 60,533 | 91,618 | **1,349,326** |
| **Usage Tercile** |  |  |  |  |  |  |
| <12.29 Avg Daily kWh  | 0.39% | 844,268 | -289,761/1,978,297 | 53,165 | 43,755 | **800,513** |
| 12.30 - 22.39 Avg Daily kWh | 0.11% | 494,323 | -752,357/1,741,003 | 59,566 | 86,731 | **407,592** |
| >22.40 Avg Daily kWh | -0.06% | -482,021 | -2,600,548/1,636,506 | 63,227 | 169,029 | **-651,050** |

*Source: Evaluation team analysis.*

*\* denotes statistical significance at the 90% level*

*Note: Values of less than 0.01% have been rounded to 0.01%*

A finding of near-zero and statistically insignificant savings is in contradiction to what would typically be expected for a residential behavior program in its second year. At the onset of this evaluation, a power analysis and benchmarking analysis were conducted to provide Eversource with background information on other digital-only behavioral programs. In this benchmarking analysis, first year savings associated with digital-only messaging were found to range from 0.25% to 0.75% for electric and 0.15% to 0.70% for gas in the first year of the program.

Savings from behavioral programs typically ramp up after the first year. Industry research suggests that participants in residential behavioral programs save between 1.2% and 2.2% of household electric usage per year. Most waves exhibit a one- or two-year ramp-up period, with savings at the ramped-up level continuing for at least the following 5 years.[[3]](#footnote-4) However, this program is not typical given the inclusion of past-HER recipients and the program’s digital-only messaging, which could be tempering the savings estimated for PY2022.

The evaluation team also looked at impacts by subgroup, including prior HER participation status and usage tercile. Impacts for each prior HER participation status as well as each usage tercile are not statistically significant. The evaluation team performed these breakouts to determine if there was a pattern in savings within each group. For example, some literature on impacts by usage band has found the largest savings come from the largest pre-use customers. No pattern was identified within either group, and the evaluation team urges using caution when comparing point estimates as each group’s savings estimates have overlapping confidence bounds and because each group’s savings are statistically insignificant.

***3.2 Massachusetts Gas Savings***

Table 2 illustrates estimated gas savings attributed to the DEI program in Massachusetts. PY2022 gas savings in Massachusetts were approximately -0.01%, amounting to -1,278 Therms across all DEI participants; however, results were not statistically significant. When including uplift, overall savings were -11,001 Therms. As described in the previous section, typical first-year gas savings for a digital-only residential behavior program can range from 0.15% to 0.70%. Given the program is now in its second year, the evaluation team would have expected savings to be at least within the range of first-year savings.

Table 2. PY2022 Massachusetts DEI Gas Savings

| Savings Category | Verified Net Savings, Prior to Uplift Adjustment |  | Savings Uplift in Other EE Programs, kWh | DEI Savings net of Uplift, kWh |
| --- | --- | --- | --- | --- |
| % | Therms | 90% Confidence Interval (Therms) | Treatment Count |
| Overall | -0.01% | -1,278 | -99,447/96,891 | 54,606 | 9,722 | **-11,001** |
| **Prior HER Status** |  |  |
| HER Recipient | 0.14% | 26,710 | -32,417/85,837 | 22,073 | 3,881 | **22,830** |
| HER Control | -0.22% | -28,926 | -76,580/18,728 | 14,509 | 2,604 | **-31,530** |
| No HER engagement | 0.01% | 1,419 | -61,424/64,262 | 18,024 | 3,211 | **-1,792** |
| **Usage Tercile** |  |  |  |  |  |  |
| <1.73 Avg Daily Therms | 0.15% | 13,689 | -32,076/59,454 | 17,126 | 1,771 | **11,919** |
| 1.74 - 2.65 Avg Daily Therms | 0.04% | 5,837 | -37,829/49,503 | 18,478 | 2,990 | **2,847** |
| >2.66 Avg Daily Therms | -0.05% | -11,381 | -81,325/58,563 | 19,002 | 4,796 | **-16,177** |

*Source: Evaluation team analysis.*

*\* denotes statistical significance at the 90% level*

*Note: Values between -0.01% and 0.00% have been rounded to -0.01%. Values between 0.00% and 0.01% have been rounded to 0.01%*

The evaluation team also looked at impacts by subgroup, including prior HER participation status and usage tercile. Impacts for each prior HER participation status as well as each usage tercile are not statistically significant. The evaluation team performed these breakouts to determine if there was a pattern in savings within each group. For example, some literature on impacts by usage band has found the largest savings come from the largest pre-use customers. No pattern was identified within either group, and the evaluation team urges using caution when comparing point estimates as each group’s savings estimates have overlapping confidence bounds and because each group’s savings are statistically insignificant.

***3.3 New Hampshire Electric Savings***

Table 3 illustrates estimated electric savings attributed to the DEI program in New Hampshire. Program year 2022 electric savings in New Hampshire were statistically significant overall at 0.27%, amounting to 2,455,408 kWh across all DEI participants. When including uplift, overall savings were 2,271,840 kWh. Savings of approximately 0.27% prior to uplift adjustment is in line with the 0.25% to 0.70% range of first-year electric savings for a digital-only program. Given this is the program’s second year, the evaluation team expected to see savings closer to the upper end (0.70%) of the cited range for a pilot year of a digital-only program.

Table . PY2022 New Hampshire DEI Electric Savings

| Savings Category | Verified Net Savings, Prior to Uplift Adjustment |  | Savings Uplift in Other EE Programs, kWh | DEI Savings net of Uplift, kWh |
| --- | --- | --- | --- | --- |
| % | kWh | 90% Confidence Interval (kWh) | Treatment Count |
| Overall | 0.27% | 2,455,408\* | 552,390/4,358,426 | 100,454 | 183,567 | **2,271,840** |
| **Prior HER Status** |  |  |
| HER Recipient | 0.63% | 981,332\* | 280,864/1,681,800 | 17,730 | 31,038 | **950,294** |
| HER Control | 0.20% | 194,767 | -357,649/747,183 | 11,178 | 19,639 | **175,128** |
| No HER engagement | 0.18% | 1,179,258 | -503,584/2,862,099 | 71,546 | 132,917 | **1,046,341** |
| **Usage Tercile** |  |  |  |  |  |  |
| <15.79 Avg Daily kWh  | 0.32% | 475,474 | -271,007/1,221,956 | 32,558 | 29,841 | **445,633** |
| 15.8 - 25.79 Avg Daily kWh | 0.49% | 1,347,210\* | 465,014/2,229,407 | 33,762 | 54,662 | **1,292,548** |
| >25.80 Avg Daily kWh | 0.07% | 361,119 | -1,122,285/1,844,524 | 34,134 | 96,959 | **264,160** |

*Source: Evaluation team analysis.*

*\* denotes statistical significance at the 90% level*

*Note: Values between -0.01% and 0.00% have been rounded to -0.01%. Values between 0.00% and 0.01% have been rounded to 0.01%*

Examining impacts based on prior HER participation status contained in Table 3, prior HER participants had statistically significant savings, while the other two subgroups had statistically insignificant savings. Since each group’s savings estimates have overlapping confidence bounds, and two of three groups’ savings are statistically insignificant, a robust and definitive assessment of what is driving savings differences between HER subgroups is not feasible.

The evaluation team also looked at impacts by usage tercile. The evaluation team performed this breakout to determine if there was a pattern in savings across the terciles. For example, some literature on impacts by usage band has found the largest savings come from the largest pre-use customers. No pattern was identified across terciles, and the evaluation team urges using caution when comparing point estimates as each group’s savings estimates have overlapping confidence bounds and because each tercile’s savings are statistically insignificant.

***3.4 Connecticut Electric Savings***

Table 4 illustrates estimated electric savings attributed to the DEI program in Connecticut. PY2022 electric savings in Connecticut were approximately 0.11%, amounting to 1,080,928 kWh across all DEI participants; however, results were not statistically significant. When including uplift, overall savings were 865,426 kWh. As described in the previous section, typical first-year savings for a digital-only residential behavior program can range from 0.15% to 0.70%. Given the program is now in its second year, the evaluation team would have expected savings to be at least within the range of first-year savings.

Table . PY2022 Connecticut DEI Electric Savings

| Savings Category | Verified Net Savings, Prior to Uplift Adjustment |  | Savings Uplift in Other EE Programs, kWh | DEI Savings net of Uplift, kWh |
| --- | --- | --- | --- | --- |
| % | kWh | 90% Confidence Interval (kWh) | Treatment Count |
| Overall | 0.11% | 1,080,928 | -815,856/2,977,713 | 107,998 | 197,350 | **865,426** |
| **Prior HER Status** |  |  |
| HER Participants | 0.01% | 40,206 | -1,516,282/ 1,596,694 | 70,286 | 142,591 | **-102,385** |
| Non-HER Participants | 0.38% | 1,040,722 | -40,829/2,122,273 | 37,712 | 54,481 | **986,241** |
| **Usage Tercile** |  |  |  |  |  |  |
| <18.9 Avg Daily kWh  | 0.28% | 494,352 | -301,681/ 1,290,385 | 36,430 | 35,441 | **458,911** |
| 18.9 – 30.5 Avg Daily kWh | 0.05% | 150,003 | -830,889/ 1,130,894 | 35,524 | 62,304 | **87,698** |
| >30.5 Avg Daily kWh | 0.07% | 344,043 | -1,104,496/ 1,792,581 | 36,044 | 101,879 | **242,163** |

*Source: Evaluation team analysis.*

*\* denotes statistical significance at the 90% level*

*Note: Values between -0.01% and 0.00% have been rounded to -0.01%. Values between 0.00% and 0.01% have been rounded to 0.01%*

The evaluation team also looked at impacts by subgroup, including prior HER participation status and usage tercile. Impacts for each prior HER participation status as well as each usage tercile are not statistically significant. The evaluation team performed these breakouts to determine if there was a pattern in savings within each group. For example, some literature on impacts by usage band has found the largest savings come from the largest pre-use customers. No pattern was identified within either group, and the evaluation team urges using caution when comparing point estimates as each group’s savings estimates have overlapping confidence bounds and because each group’s savings are statistically insignificant.

***3.5 Connecticut Gas Savings***

Table 5 illustrates estimated gas savings attributed to the DEI program in Connecticut. PY2022 gas savings in Connecticut were approximately 0.28%, amounting to 37,105 Therms across all DEI participants; however, results were not statistically significant. When including uplift, overall savings were 33,605 Therms. While savings of approximately 0.28% prior to uplift adjustment is in-line with the 0.15% to 0.70% range of first-year gas savings for a digital-only program, these savings were statistically insignificant. Additionally, since this is the program’s second year, the evaluation team expected to see savings closer to the upper end (0.70%) of the cited range for a pilot year of a digital-only program.

Table . PY 2022 Connecticut DEI Gas Savings

| Savings Category | Verified Net Savings, Prior to Uplift Adjustment |  | Savings Uplift in Other EE Programs, kWh | DEI Savings net of Uplift, kWh |
| --- | --- | --- | --- | --- |
| % | Therms | 90% Confidence Interval (Therms) | Treatment Count |
| Overall | 0.28% | 37,105 | -11,296/ 85,505 | 18,591 | 2,633 | **33,605** |
| **Prior HER Status** |  |  |
| HER Participants | 0.30% | 25,593 | -9,808/ 60,994 |  11,066  | 1,696 | **23,897** |
| Non-HER Participants | 0.25% | 11,512 | -21,680/ 44,704 |  7,525  | 933 | **10,579** |
| **Usage Tercile** |  |  |  |  |  |  |
| <1.55 Avg Daily Therms  | -0.36% | -9,040 | -33,454/ 15,374 |  6,232  | 496 | **-9,535** |
| 1.55 – 2.40 Avg Daily Therms | 0.45% | 19,219 | -5,013/ 43,452 |  6,090  | 854 | **18,365** |
| >2.40 Avg Daily Therms | 0.45% | 29,746 | -5,610/65,102 |  6,269  | 1,331 | **28,416** |

*Source: Evaluation team analysis.*

*\* denotes statistical significance at the 90% level*

*Note: Values between -0.01% and 0.00% have been rounded to -0.01%. Values between 0.00% and 0.01% have been rounded to 0.01%*

The evaluation team also looked at impacts by subgroup, including prior HER participation status and usage tercile. Impacts for each prior HER participation status as well as each usage tercile are not statistically significant. The evaluation team performed these breakouts to determine if there was a pattern in savings within each group. For example, some literature on impacts by usage band has found the largest savings come from the largest pre-use customers. No pattern was identified within either group, and the evaluation team urges using caution when comparing point estimates as each group’s savings estimates have overlapping confidence bounds and because each group’s savings are statistically insignificant.

# 4. Discussion

In this memorandum, the evaluation team provided the second-year evaluation results for the DEI program. The evaluation results provided definitive electric impact estimates for New Hampshire. However, for Connecticut and Massachusetts, the evaluation results were not definitive. To further understand the results for Connecticut and Massachusetts, the evaluation team investigated savings for customers who were and were not previously in the discontinued HER program, as well as for DEI participants in the lower, middle, and upper 33% of energy users. However, these investigations did not reveal any statistically valid patterns and provided limited insights into the uncertainty of results for Connecticut and Massachusetts.

While the results from this evaluation do not provide any clear explanations for results for Connecticut and Massachusetts, the evaluation team and stakeholders discussed several potential hypotheses that Eversource could consider for future research or program design changes. In the paragraphs below, the evaluation team provides a discussion of these potential hypotheses separated by those that are within vs outside the DEI program’s control.

Some potential hypotheses for the results in this evaluation include trends outside the DEI program’s control. For example, customers may have become fatigued by energy usage messaging, and as a result, tune the messaging out. As behavioral programs have expanded across the US, utilities and energy efficiency programs have provided more messaging about energy usage and energy savings. This expanded messaging includes new online platforms that enable customers to review their energy usage and bills online or on their phones as well as an increase in customers reviewing their energy bills via e-mail. Another trend outside the DEI program’s control is saturation in the behavioral energy efficiency space. After years of increased messaging with few-if-any new energy savings tips, customers may already be implementing the behaviors that these programs encourage. Lastly, the increasing efficiency of equipment (including baseline equipment) may be reducing the opportunities for behavioral savings. For example, LED lighting uses less energy, and its widespread adoption may have reduced the impact of what can be expected from turning off the lights.

Stakeholders also discussed potential hypotheses for the results in this evaluation that include topics within the DEI program’s control. For example, stakeholders discussed whether the open rates are different in New Hampshire compared to both Connecticut and Massachusetts. The DEI program team can potentially conduct internal research to answer that question, as well as conduct research using quasi-experimental methods to determine if customers with higher open rates save more energy than other customers. If the DEI program team determines that open rates strongly correlate with savings, they can use open rates as a proxy for savings in other internal research. Also, stakeholders discussed some of the differences between the discontinued Oracle program and the DEI program. These differences offer opportunities for the DEI program to potentially conduct internal testing or research, such as using AB testing or other similar methods, to determine if changes in the DEI program design can provide better savings. The differences include: sending digital only vs digital and paper messaging, the frequency and timing of the messages, whether to include peer comparisons, and the design of the report.

Finally, stakeholders also noted that these issues are not unique to Eversource and there is currently research in other jurisdictions seeking to answer similar questions. Further considerations on how to study these issues are discussed in Section 5.

# 5. Key Findings & Considerations

The evaluation team provided savings estimates by state and fuel type in the previous section. In addition, the evaluation team provided savings estimates for subgroups of Eversource customers based on their prior participation status in Eversource’s discontinued HER program, and for the lower, middle, and upper 33% of energy users.

Based on the design of this study, overall savings provided in the tables presented above are distinct and appropriate for Eversource to add together by state and fuel type in their filings when warranted. The evaluation team notes that DEI savings were only statistically significant for New Hampshire electric participants. Savings estimates for gas and electric participants in Connecticut and Massachusetts were statistically insignificant and the evaluation team encourages stakeholders in each state to determine the appropriate course of action for claiming savings. For example, stakeholders could decide to (a) accept the results as the best estimate of savings, or (b) determine that the uncertainty is too high and savings should be considered zero. Similarly, given the uncertainty in the results, stakeholders may decide to exclude the uplift adjustment for any savings claimed.

***5.1 Key Findings***

* **Delivered Energy Insights yielded statistically significant electric savings for New Hampshire.** In a 2020 benchmarking analysis on digital-only behavioral programs, the evaluation team found that first year savings associated with digital-only messaging ranged from 0.25% to 0.75% for electricity and from 0.15% to 0.70% for gas. The second-year savings results for New Hampshire fall within this range of typical first year savings (0.27%).
* **Delivered Energy Insights continued to yield savings that are not statistically different than zero in its second year for Connecticut and Massachusetts.** In a 2020 benchmarking analysis on digital-only behavioral programs, the evaluation team found that first year savings associated with digital-only messaging ranged from 0.25% to 0.75% for electricity and from 0.15% to 0.70% for gas. Based on evaluation results, second-year savings, which would ideally exceed those observed in the first year, fell short of the first-year benchmark range for both Connecticut and New Hampshire.
* **There is not a clear directional pattern between prior Home Energy Reports participation status and savings in the Delivered Energy Insights program.** For some states and fuel types, DEI electric participants with no prior HER engagement had higher savings than both previous HER recipients and controls. For other states and fuel types, the trend was the reverse; however, the New Hampshire HER recipient group was the only statistically significant result in the analysis of savings based on prior HER participation status. All other subgroups had statistically insignificant savings and confidence bounds that overlapped between groups. As such, no definitive directional pattern (e.g., prior HER participants tend to save the most) could be ascertained from the results.
* **There is not a clear directional pattern related to usage terciles**. Of all usage terciles examined, only one tercile (New Hampshire, middle 33% of energy users) yielded statistically significant savings. For all other terciles, results were statistically insignificant, and all terciles’ savings estimates had overlapping confidence bounds. Given the lack of statistical significance as well as the overlapping confidence bounds, no definitive directional patterns can be ascertained from these results.

***5.2 Considerations***

* **Consider conducting participant messaging research, including an assessment of recall rates and customer comprehension of the messaging.** The DEI program continued to yield savings that are not statistically different than zero in its second year for Connecticut and Massachusetts. Limited impacts may be attributed to treatment customers not understanding or engaging with program messaging. Research into participant engagement could help identify program shortcomings that should be addressed.

# Appendix A – Equivalency Checks

**A.1 Equivalency Check Methodology**

The DEI behavioral program was launched as a randomized control trial (RCT), in which individual customers were randomly assigned to the treatment group (i.e., they receive DEI emails) or the control group (i.e., they do not receive DEI emails) for estimating changes in electric and gas consumption due to the program. Since the program randomly assigned customers to the treatment and control groups, pre-treatment energy use should be equivalent between the groups. To ensure the validity of savings estimates derived from this evaluation, it is important to confirm that this assumption holds.

The evaluation team conducted randomization checks on each DEI analysis group as originally assigned (or reassigned) and with opt-outs, move-outs, and any ineligible customers removed. For each DEI analysis group, the evaluation team used four methods of determining the equivalency of treatment and control groups:

* T-tests on monthly differences in mean energy use between treatment and control groups in each of the 12 months leading up to program treatment. A significant difference (p < 0.10) indicates that pre-period usage is dissimilar between groups.
* Regression analysis of pre-period usage in the 12 months leading up to program treatment, with treatment/control identification as a covariate. A significant effect (p < 0.10) of the group category indicates that pre-period usage is dissimilar between groups.
* Plot of average daily energy (i.e., electric and gas) consumption per month between treatment and controlto visually assess the magnitude of any pre-period differences in energy consumption.
* Inspection of median family incomes and median individual incomes of ZIP codes in which treatment and control customers reside. For this analysis, the evaluation team tabulated the number and percentage of treatment and control customers whose ZIP code fell within specified income ranges.

**A.2 Equivalency Check Results**

The evaluation team has provided detailed tables containing the results of randomization checks on treatment and control customers. Across all randomization checks, the evaluation team found no marked differences between the treatment and control customers. Below, the evaluation team provides more detailed findings from randomization checks conducted.

* T-tests on monthly differences in mean energy use between treatment and control groups in each of the 12 months leading up to program treatment indicated no significant differences in usage between treatment and control groups in Connecticut or Massachusetts. In New Hampshire, one month had a statistically significant difference in usage (January 2020), but this is likely a function of a limited number of customer bills received for that month. All subsequent months of the pre-period had no statistically significant differences in energy use between treatment and control. Table 6 through Table 12 provide T-test results validating the integrity of the RCT for the pre-period of each jurisdiction.
* Regression analysis of pre-period usage in the 12 months leading up to program treatment indicated no significant difference between treatment and control. Table 13 provides the results of the regression based RCT check for each jurisdiction. All jurisdictions evaluated passed this RCT check as well.
* Inspection of median family incomes and median individual incomes of ZIP codes in which treatment and control customers reside revealed no marked differences in incomes represented by the two groups. Table 14 through Table 19 provide the results of the additional demographic balance checks based on family and individual median income. There is balance between treatment and control customers on income in MA, NH, and CT with the bulk of participants in zip codes with incomes close to each state’s median.

***A.2.1 Monthly T-Test Balance Check Results***

Table . Massachusetts Electric T-Test RCT Check

| Month | Mean Diff | Mean Treatment | Mean Control | T-stat | P value | Degrees of Freedom | Confidence Low | Confidence High |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 01/2020 | 0.03 | 19.95 | 19.92 | 0.51 | 0.61 | 410748 | -0.06 | 0.11 |
| 02/2020 | 0.03 | 18.86 | 18.83 | 0.51 | 0.61 | 375342 | -0.06 | 0.12 |
| 03/2020 | -0.02 | 18.10 | 18.12 | -0.34 | 0.74 | 235352 | -0.12 | 0.08 |
| 04/2020 | 0.01 | 17.49 | 17.47 | 0.27 | 0.79 | 334292 | -0.06 | 0.09 |
| 05/2020 | 0.01 | 16.52 | 16.51 | 0.21 | 0.83 | 283187 | -0.07 | 0.09 |
| 06/2020 | -0.01 | 18.47 | 18.48 | -0.28 | 0.78 | 368584 | -0.09 | 0.06 |
| 07/2020 | -0.01 | 25.47 | 25.48 | -0.13 | 0.90 | 320999 | -0.12 | 0.10 |
| 08/2020 | 0.02 | 28.69 | 28.67 | 0.27 | 0.78 | 366037 | -0.09 | 0.13 |
| 09/2020 | 0.00 | 20.66 | 20.66 | 0.04 | 0.97 | 393562 | -0.08 | 0.08 |
| 10/2020 | -0.01 | 16.45 | 16.46 | -0.31 | 0.76 | 255984 | -0.09 | 0.06 |
| 11/2020 | -0.01 | 16.78 | 16.79 | -0.16 | 0.87 | 341634 | -0.07 | 0.06 |
| 12/2020 | 0.04 | 19.23 | 19.20 | 0.72 | 0.47 | 356943 | -0.05 | 0.12 |

*Source: Evaluation team analysis.*

*Note: results are from a Welch Two-Sample T-test*

Table . Massachusetts Gas T-Test RCT Check

| Month | Mean Diff | Mean Treatment | Mean Control | T-stat | P value | Degrees of Freedom | Confidence Low | Confidence High |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 01/2020 | 0.01 | 4.64 | 4.63 | 0.45 | 0.65 | 123546 | -0.02 | 0.03 |
| 02/2020 | 0.01 | 4.78 | 4.77 | 0.68 | 0.50 | 102304 | -0.01 | 0.03 |
| 03/2020 | 0.00 | 3.98 | 3.98 | 0.16 | 0.87 | 56012 | -0.03 | 0.03 |
| 04/2020 | 0.01 | 3.12 | 3.11 | 0.90 | 0.37 | 110023 | -0.01 | 0.03 |
| 05/2020 | 0.00 | 2.39 | 2.39 | -0.06 | 0.96 | 89660 | -0.02 | 0.02 |
| 06/2020 | 0.00 | 0.95 | 0.95 | 0.14 | 0.89 | 107308 | -0.01 | 0.01 |
| 07/2020 | 0.00 | 0.61 | 0.61 | -0.35 | 0.73 | 91914 | -0.01 | 0.01 |
| 08/2020 | 0.00 | 0.53 | 0.52 | 0.43 | 0.67 | 98073 | 0.00 | 0.01 |
| 09/2020 | 0.00 | 0.59 | 0.59 | 0.73 | 0.47 | 110466 | 0.00 | 0.01 |
| 10/2020 | 0.00 | 0.91 | 0.91 | -0.38 | 0.71 | 81987 | -0.01 | 0.01 |
| 11/2020 | 0.00 | 2.04 | 2.03 | 0.19 | 0.85 | 103901 | -0.01 | 0.01 |
| 12/2020 | 0.00 | 3.40 | 3.40 | -0.35 | 0.73 | 116373 | -0.02 | 0.01 |

*Source: Evaluation team analysis.*

*Note: results are from a Welch Two-Sample T-test*

Table . New Hampshire Electric T-Test RCT Check

| Month | Mean Diff | Mean Treatment | Mean Control | T-stat | P value | Degrees of Freedom | Confidence Low | Confidence High |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 01/2020 | -0.61 | 22.52 | 23.13 | -2.01 | 0.04 | 13739 | -1.11 | -0.11 |
| 02/2020 | 0.03 | 22.86 | 22.83 | 0.38 | 0.70 | 199055 | -0.10 | 0.15 |
| 03/2020 | -0.02 | 21.48 | 21.50 | -0.29 | 0.77 | 212516 | -0.13 | 0.09 |
| 04/2020 | 0.02 | 21.05 | 21.04 | 0.25 | 0.81 | 199305 | -0.09 | 0.12 |
| 05/2020 | -0.02 | 19.97 | 19.99 | -0.34 | 0.73 | 212800 | -0.11 | 0.07 |
| 06/2020 | 0.02 | 21.92 | 21.91 | 0.25 | 0.80 | 212816 | -0.09 | 0.12 |
| 07/2020 | 0.02 | 27.94 | 27.92 | 0.27 | 0.79 | 212810 | -0.11 | 0.15 |
| 08/2020 | -0.02 | 30.82 | 30.84 | -0.18 | 0.86 | 212813 | -0.15 | 0.12 |
| 09/2020 | -0.01 | 23.28 | 23.29 | -0.12 | 0.90 | 212826 | -0.11 | 0.10 |
| 10/2020 | -0.01 | 19.25 | 19.26 | -0.21 | 0.83 | 212845 | -0.10 | 0.07 |
| 11/2020 | 0.03 | 20.25 | 20.23 | 0.43 | 0.67 | 192342 | -0.07 | 0.12 |
| 12/2020 | 0.01 | 23.17 | 23.16 | 0.11 | 0.91 | 212731 | -0.11 | 0.12 |

*Source: Evaluation team analysis.*

*Note: results are from a Welch Two-Sample T-test*

Table . Connecticut Electric T-Test RCT Check, Non-HER Customers

| Month | Mean Diff | Mean Treatment | Mean Control | P value |
| --- | --- | --- | --- | --- |
| 01/2020 | 0.16 | 20.62 | 20.45 | 0.22 |
| 02/2020 | 0.18 | 19.85 | 19.68 | 0.19 |
| 03/2020 | 0.07 | 18.14 | 18.08 | 0.56 |
| 04/2020 | 0.08 | 17.78 | 17.70 | 0.46 |
| 05/2020 | 0.03 | 16.81 | 16.78 | 0.72 |
| 06/2020 | -0.01 | 18.20 | 18.21 | 0.93 |
| 07/2020 | -0.08 | 25.34 | 25.42 | 0.54 |
| 08/2020 | -0.10 | 27.55 | 27.65 | 0.46 |
| 09/2020 | -0.07 | 21.13 | 21.20 | 0.51 |
| 10/2020 | 0.00 | 16.23 | 16.23 | 0.96 |
| 11/2020 | 0.00 | 16.97 | 16.97 | 0.96 |
| 12/2020 | 0.05 | 20.17 | 20.13 | 0.70 |

*Source: Evaluation team analysis.*

Table . Connecticut Electric T-Test RCT Check, Prior HER Customers

| Month | Mean Diff | Mean Treatment | Mean Control | P value |
| --- | --- | --- | --- | --- |
| 01/2020 | 0.04 | 30.16 | 30.13 | 0.73 |
| 02/2020 | 0.06 | 28.52 | 28.47 | 0.56 |
| 03/2020 | 0.03 | 26.26 | 26.23 | 0.75 |
| 04/2020 | 0.02 | 25.85 | 25.83 | 0.76 |
| 05/2020 | 0.00 | 24.53 | 24.54 | 0.96 |
| 06/2020 | -0.04 | 27.18 | 27.23 | 0.62 |
| 07/2020 | 0.04 | 37.74 | 37.69 | 0.69 |
| 08/2020 | 0.07 | 39.90 | 39.83 | 0.53 |
| 09/2020 | 0.05 | 30.86 | 30.80 | 0.57 |
| 10/2020 | 0.08 | 23.37 | 23.29 | 0.24 |
| 11/2020 | 0.06 | 24.18 | 24.13 | 0.40 |
| 12/2020 | 0.07 | 28.90 | 28.83 | 0.46 |

*Source: Evaluation team analysis.*

Table . Connecticut Gas T-Test RCT Check, Non-HER Customers

| Month | Mean Diff | Mean Treatment | Mean Control | P value |
| --- | --- | --- | --- | --- |
| 01/2020 | -0.04 | 3.52 | 3.56 | 0.25 |
| 02/2020 | -0.04 | 3.59 | 3.63 | 0.27 |
| 03/2020 | -0.02 | 3.01 | 3.03 | 0.42 |
| 04/2020 | -0.02 | 2.22 | 2.24 | 0.43 |
| 05/2020 | -0.01 | 1.68 | 1.69 | 0.67 |
| 06/2020 | 0.00 | 0.72 | 0.72 | 0.81 |
| 07/2020 | 0.00 | 0.48 | 0.48 | 0.65 |
| 08/2020 | 0.00 | 0.44 | 0.44 | 0.96 |
| 09/2020 | 0.00 | 0.46 | 0.46 | 0.90 |
| 10/2020 | 0.00 | 0.67 | 0.67 | 0.67 |
| 11/2020 | -0.02 | 1.56 | 1.58 | 0.19 |
| 12/2020 | -0.03 | 2.68 | 2.71 | 0.33 |

*Source: Evaluation team analysis.*

Table . Connecticut Gas T-Test RCT Check, Prior HER Customers

| Month | Mean Diff | Mean Treatment | Mean Control | P value |
| --- | --- | --- | --- | --- |
| 01/2020 | -0.02 | 4.31 | 4.33 | 0.56 |
| 02/2020 | -0.01 | 4.40 | 4.41 | 0.78 |
| 03/2020 | 0.00 | 3.69 | 3.69 | 0.89 |
| 04/2020 | 0.00 | 2.72 | 2.72 | 0.91 |
| 05/2020 | 0.00 | 2.06 | 2.06 | 0.87 |
| 06/2020 | 0.00 | 0.89 | 0.89 | 0.64 |
| 07/2020 | 0.00 | 0.60 | 0.59 | 0.43 |
| 08/2020 | 0.00 | 0.55 | 0.55 | 0.62 |
| 09/2020 | 0.00 | 0.57 | 0.57 | 0.69 |
| 10/2020 | 0.00 | 0.81 | 0.80 | 0.81 |
| 11/2020 | -0.01 | 1.88 | 1.89 | 0.53 |
| 12/2020 | -0.01 | 3.24 | 3.25 | 0.66 |

*Source: Evaluation team analysis.*

***A.2.2 Pre-Period Regression Balance Check Results***

Table . Regression RCT Check

| State | Fuel | Treatment Estimate | Standard Error | T Statistic | P value |
| --- | --- | --- | --- | --- | --- |
| MA | Electric | 0.0075 | 0.0426 | 0.1772 | 0.8593 |
| NH | Electric | -0.0004 | 0.0584 | -0.0070 | 0.9944 |
| CT | Electric | 0.04 | 0.06 | 0.67 | 0.50 |
| MA | Therms | 0.0024 | 0.0068 | 0.3491 | 0.7270 |
| CT | Therms | -0.01 | 0.01 | -0.65 | 0.52 |

*Source: Evaluation team analysis.*

***A.2.3 Income Balance Check Results***

Table . Massachusetts Family Income Balance Check

| State | Median Family Income | Control | Treatment | Percent of Control | Percent of Treatment |
| --- | --- | --- | --- | --- | --- |
| MA |  <29,999  |  971  |  1,034  | 0% | 0% |
| MA |  30,000-39,999  |  4,219  |  4,152  | 2% | 2% |
| MA |  40,000-49,999  |  13,680  |  13,553  | 5% | 5% |
| MA |  50,000-59,999  |  6,793  |  7,029  | 2% | 3% |
| MA |  60,000-69,999  |  14,219  |  14,245  | 5% | 5% |
| MA |  70,000-79,999  |  26,539  |  26,615  | 10% | 10% |
| MA |  80,000-89,999  |  15,247  |  15,061  | 6% | 6% |
| MA |  90,000-99,999  |  15,939  |  15,935  | 6% | 6% |
| MA |  100,000-109,999  |  16,545  |  16,336  | 6% | 6% |
| MA |  110,000-119,999  |  21,870  |  22,081  | 8% | 8% |
| MA | 120,000-129,999  |  17,733  |  17,645  | 7% | 6% |
| MA |  130,000-139,999  |  15,826  |  15,720  | 6% | 6% |
| MA |  140,000-149,999  |  25,299  |  25,157  | 9% | 9% |
| MA |  150,000-159,999  |  23,149  |  23,155  | 9% | 9% |
| MA |  160,000-169,999  |  7,788  |  7,973  | 3% | 3% |
| MA |  170,000-179,999  |  5,942  |  6,058  | 2% | 2% |
| MA |  180,000-189,999  |  7,729  |  7,761  | 3% | 3% |
| MA |  190,000-199,999  |  2,982  |  3,089  | 1% | 1% |
| MA |  200,000-209,999  |  3,462  |  3,412  | 1% | 1% |
| MA |  210,000-219,999  |  9,966  |  9,926  | 4% | 4% |
| MA |  220,000-229,999  |  6,739  |  6,732  | 2% | 2% |
| MA |  230,000-239,999  |  3,570  |  3,539  | 1% | 1% |
| MA |  >240,000 |  4,798  |  4,825  | 2% | 2% |
| MA | -  |  997  |  967  | 0% | 0% |

*Source: Evaluation team analysis of 2021 US Census American Community Survey data.*

Table . Massachusetts Individual Income Balance Check

| State | Median Individual Income | Control | Treatment | Percent of Control | Percent of Treatment |
| --- | --- | --- | --- | --- | --- |
| MA | <9,999 | 1068 | 1,100 | 0% | 0% |
| MA | 10,000-19,000 | 8,470 | 8,302 | 3% | 3% |
| MA | 20,000-29,999 | 22,731 | 22,930 | 8% | 8% |
| MA |  30,000-39,999  | 50,928 | 50,956 | 19% | 19% |
| MA |  40,000-49,999  | 46,401 | 46,110 | 17% | 17% |
| MA |  50,000-59,999  | 57,185 | 57,324 | 21% | 21% |
| MA |  60,000-69,999  | 16,074 | 15,949 | 6% | 6% |
| MA |  70,000-79,999  | 26,812 | 26,708 | 10% | 10% |
| MA |  80,000-89,999  | 13,481 | 13,607 | 5% | 5% |
| MA |  90,000-99,999  | 14,296 | 14,285 | 5% | 5% |
| MA |  100,000-109,999  | 5,167 | 5,185 | 2% | 2% |
| MA |  110,000-119,999  | 5,377 | 5,549 | 2% | 2% |
| MA | 120,000-129,999  | 470 | 444 | 0% | 0% |
| MA |  130,000-139,999  | 25 | 24 | 0% | 0% |
| MA | >140,000 | 765 | 708 | 0% | 0% |
| MA | - | 2,752 | 2,819 | 1% | 1% |

*Source: Evaluation team analysis of 2021 US Census American Community Survey data.*

Table . New Hampshire Family Income Balance Check

| State | Median Family Income | Control | Treatment | Percent of Control | Percent of Treatment |
| --- | --- | --- | --- | --- | --- |
| NH |  <39,999  |  28  |  18  | 0% | 0% |
| NH |  40,000-49,999  |  148  |  131  | 0% | 0% |
| NH |  50,000-59,999  |  2,941  |  2,876  | 3% | 3% |
| NH |  60,000-69,999  |  7,719  |  7,910  | 7% | 7% |
| NH |  70,000-79,999  |  17,544  |  17,594  | 16% | 16% |
| NH |  80,000-89,999  |  10,719  |  10,955  | 10% | 10% |
| NH |  90,000-99,999  |  15,877  |  15,856  | 15% | 15% |
| NH |  100,000-109,999  |  18,333  |  17,935  | 17% | 17% |
| NH |  110,000-119,999  |  8,994  |  9,018  | 8% | 8% |
| NH | 120,000-129,999  |  11,452  |  11,488  | 11% | 11% |
| NH |  130,000-139,999  |  5,845  |  5,807  | 5% | 5% |
| NH |  140,000-149,999  |  1,526  |  1,409  | 1% | 1% |
| NH |  150,000-159,999  |  2,469  |  2,465  | 2% | 2% |
| NH |  160,000-169,999  |  4,317  |  4,482  | 4% | 4% |
| NH |  170,000-179,999  |  194  |  187  | 0% | 0% |
| NH |  180,000-189,999  |  4  |  2  | 0% | 0% |
| NH |  >210,000  |  112  |  108  | 0% | 0% |
| NH | - |  311  |  294  | 0% | 0% |

*Source: Evaluation team analysis of 2021 US Census American Community Survey data.*

Table . New Hampshire Individual Income Balance Check

| State | Median Individual Income | Control | Treatment | Percent of Control | Percent of Treatment |
| --- | --- | --- | --- | --- | --- |
| NH | <9,999 |  4  |  6  | 0% | 0% |
| NH | 10,000-19,000 |  2,146  |  2,103  | 2% | 2% |
| NH | 20,000-29,999 |  10,122  |  10,289  | 9% | 9% |
| NH |  30,000-39,999  |  28,035  |  28,251  | 26% | 26% |
| NH |  40,000-49,999  |  27,679  |  27,633  | 26% | 25% |
| NH |  50,000-59,999  |  19,348  |  19,201  | 18% | 18% |
| NH |  60,000-69,999  |  13,719  |  13,797  | 13% | 13% |
| NH |  70,000-79,999  |  5,004  |  4,917  | 5% | 5% |
| NH |  80,000-89,999  |  493  |  454  | 0% | 0% |
| NH |  >90,000  |  707  |  611  | 1% | 1% |
| NH | - |  1,276  |  1,273  | 1% | 1% |

*Source: Evaluation team analysis of 2021 US Census American Community Survey data.*

Table . Connecticut Electric Family Income Balance Check

| State | Median Family Income | Control | Treatment | Percent of Control | Percent of Treatment |
| --- | --- | --- | --- | --- | --- |
| CT | 10,000-19,000 | 36 | 56 | 0% | 0% |
| CT | 20,000-29,999 | 337 | 312 | 0% | 0% |
| CT |  30,000-39,999  | 2,094 | 2,047 | 2% | 2% |
| CT |  40,000-49,999  | 5,957 | 5,927 | 5% | 5% |
| CT |  50,000-59,999  | 5,698 | 5,679 | 5% | 5% |
| CT |  60,000-69,999  | 9,867 | 9,867 | 8% | 8% |
| CT |  70,000-79,999  | 12,222 | 12,260 | 10% | 10% |
| CT |  80,000-89,999  | 17,159 | 17,132 | 14% | 14% |
| CT |  90,000-99,999  | 18,826 | 18,683 | 15% | 15% |
| CT |  >100,000  | 52,472 | 52,739 | 42% | 42% |
| CT | - | 373 | 340 | 0% | 0% |

*Source: Evaluation team analysis of 2021 US Census American Community Survey data.*

Table . Connecticut Gas Family Income Balance Check

| State | Median Family Income | Control | Treatment | Percent of Control | Percent of Treatment |
| --- | --- | --- | --- | --- | --- |
| CT | 10,000-19,000 | 33 | 30 | 0% | 0% |
| CT | 20,000-29,999 | 1 | - | 0% | 0% |
| CT |  30,000-39,999  | 0 | 0 | 0% | 0% |
| CT |  40,000-49,999  | 998 | 1,100 | 4% | 5% |
| CT |  50,000-59,999  | 2,057 | 2,085 | 9% | 9% |
| CT |  60,000-69,999  | 5,104 | 5,166 | 23% | 23% |
| CT |  70,000-79,999  | 2,494 | 2,493 | 11% | 11% |
| CT |  80,000-89,999  | 4,322 | 4,213 | 19% | 19% |
| CT |  90,000-99,999  | 2,239 | 2,323 | 10% | 10% |
| CT |  >100,000  | 5,183 | 5,010 | 23% | 22% |
| CT | - | 15 | 21 | 0% | 0% |

*Source: Evaluation team analysis of 2021 US Census American Community Survey data.*

# Appendix B – Impact Evaluation Methodology

**B.1 Data Cleaning**

Certain customers and data points were removed from the analysis over several steps:

1. Observations missing bill dates that could not be backed out from additional information.
2. Observations with missing or extreme bill duration days.
3. Observations with missing or outlier usage values. Outliers were defined separately for gas and electric usage. For gas, outliers are defined as any observation with negative average daily use or based on the magnitude of deviation above the median usage. For electric, outliers are defined based on the magnitude of deviation above and below the median usage

For each jurisdiction and fuel type, these steps removed observations and customers evenly across participants and controls. This suggests that non-random biases were not introduced during data cleaning.

**B.2 Missing Pre-Period Data**

Some accounts were missing pre-period data. Incomplete (fewer than 12 months) pre-period data can introduce bias when modeling savings. This can occur particularly when pre-period data is missing entire seasons, as pre-period usage is no longer an accurate representation of usage that may be expected within a given calendar year. In addition, due to the nature of the lagged dependent variable (LDV) model, participants with incomplete pre-period data would be dropped from the modeling process.[[4]](#footnote-5)

The evaluation team addressed missing pre-period data by imputation or by conducting the analysis at multiple levels of missing pre-period data. To implement imputation, the evaluation team replaced the missing data with substituted values (i.e., imputed data) representing average daily consumption for customers that have pre-period data. The team performed imputations independently for each subgroup (state and fuel; state, fuel, and prior HER status; state, fuel, and usage tercile) and customer group (treatment and control). To conduct the analysis at multiple levels of missing pre-period data, the team ran the analysis models using different requirements for the number of pre-period observations (e.g., 3, 9, and 12 months of data). Alignment between these results indicates that data availability does not drive the estimated savings.

**B.3 Regression Approach**

To evaluate electric and gas savings attributed to the DEI program, the evaluation team estimated an LDV model for each electric and gas wave independently. The LDV model combines cross-sectional and time series data in a panel dataset. This model uses only the post-program data, with lagged energy use for the same calendar month of the pre-program period acting as a control for small systematic differences between the participant and control customers. In particular, energy use in calendar month *t* of the post-program period is a function of program participation and energy use in the same calendar month of the pre-program period. The underlying logic is that systematic differences between participants and controls will be reflected in differences in their past energy use, which is highly correlated with their current energy use. The model is detailed below.

Equation B‑1. LDV Model

$$ADC\_{kt}= β\_{1}Treatment\_{k}+ \sum\_{j}^{}β\_{2j}Month\_{jt}+\sum\_{j}^{}β\_{3j}Month\_{jt} · ADClag\_{kt}+ ε\_{kt}$$

where

***ADCkt*** = The average daily usage in kWh or Therms for customer *k* during billing cycle *t*. This is the dependent variable in the model.

***Treatmentk*** = A binary variable indicating whether customer *k* is in the treatment group (taking a value of 1) or in the control group (taking a value of 0).

***ADClagkt***= Customer *k*’s energy use in the same calendar month of the pre-program year as the calendar month of month *t*.

***Monthjt*** = A binary variable taking a value of 1 when *j* = *t* and 0 otherwise.

$ε\_{kt}$ = The cluster-robust error term for customer *k* during billing cycle *t.* Cluster-robust errors account for heteroscedasticity and autocorrelation at the customer level.[[5]](#footnote-6)

In this model, $β\_{1}$ is the estimate of average daily energy savings due to the program. Total program savings are the product of the average daily savings estimate and the total number of participant days in the analysis.

# Appendix C – Benchmarking Analysis

Please find the benchmarking analysis attached as a companion spreadsheet to this report.

1. This group is separating customers who were previously engaged in HER reports from those who never engaged in the program. For these customers, DEI is their first energy insights behavioral program. [↑](#footnote-ref-2)
2. Participant days for the annual savings calculation encompass all participants, including those without email send dates. Customers missing their first email send dates are assumed to have participated in the full 2022 program year. [↑](#footnote-ref-3)
3. Sussman, R., and M. Chikumbo (2016). “Behavior Change Programs: Status and Impact.” American Council for an Energy-Efficient Economy. <https://aceee.org/sites/default/files/publications/researchreports/b1601.pdf> [↑](#footnote-ref-4)
4. Because this model includes pre-period information as explanatory variables, if a customer is missing billing data for a certain pre-period month, then the model will drop this calendar month in the analysis period for the customer. [↑](#footnote-ref-5)
5. Ordinary Least Squares (OLS) regression models assume that the data is homoscedastic and not autocorrelated. If either of these assumptions is violated, the resulting standard errors of the parameter estimates are incorrect (usually underestimated). A random variable is heteroscedastic when the variance is not constant. A random variable is autocorrelated when the error term in one period is correlated with the error terms in at least some of the previous periods. [↑](#footnote-ref-6)