

“Lost in Translation: successes and setbacks of expanding language access to program evaluation”

Ignacio Obejero-Paz, Eversource
Kiersten Williams, Eversource

ABSTRACT

The Massachusetts Program Administrators (“PAs”) have established a foundational principle to better serve customers with historically low levels of program participation. The statewide Evaluation Management Committee has devoted time and resources to help the PAs better understand the outcomes and lessons of their efforts to help customers overcome barriers to participation. Evaluation work itself serves as a critical touchpoint with customers from different sectors to receive feedback on a range of programs and initiatives, and so the evaluation committee decided to align their strategy with the PAs’ plan, ensuring evaluation studies consider the potential of learning about segments of the population with low participation rates. The result was a new framework for how to incorporate equity-related research questions into evaluation study planning and implementation.

This paper will focus on lessons learned from the evaluation committee’s efforts to better understand one specific subset of these customers: those that speak languages other than English. Key components of the framework and how it is applied to different stages of evaluation work when learning about these customers will be described. The examples will come from work directly related to the PAs’ efforts to better serve these customers, as well as from studies where data was collected on non-English speaking ratepayers. The paper will conclude by highlighting some findings from these initial efforts, as well as the challenges that could result from applying an equity lens to evaluation studies.

Introduction

Since 2008, the Massachusetts Program Administrators (“PAs”) have submitted a detailed plan aimed at pursuing all-available and cost-effective forms of energy efficiency to regulators on a three-year cadence. Each filing has included a section on the evaluation, measurement, and verification (“EM&V”) functions and priorities intended to be conducted in the respective three-year period. During each term, EM&V activities have primarily focused on providing continuous PA program improvements and innovations, as well as accountability by ensuring claimed program impacts are accurate and credible.

For the 2025-2027 term (“the current term”), the PAs proposed to invest approximately two billion dollars for energy efficiency and electrification efforts to better serve in an equitable manner customers that have had historically low levels of program participation. The Evaluation Management Committee (“EMC”), the statewide steering committee that manages EM&V activities in Massachusetts, laid out in their 2025-2027 Strategic Evaluation Plan (“SEP”) the intent to evaluate these PA efforts and initiatives. As EM&V work requires direct interaction with customers, the EMC needed to evolve how it collects information from program participants and also non-participants to ensure it reached customers with historically low participation as well.

While the EMC has previously done work aimed at better learning about customers with low program participation rates, it was conducted within the confines of specific studies focused on better understanding participation barriers and specific customer segments. However, the

2025-2027 term was the first time the EMC deployed an equity lens across the entire EM&V portfolio. As a result, the EM&V Equity Framework (“the equity framework” or “the framework”) had to be established by the EMC to provide guidance on how the issue should be addressed across the evaluation study process.

This paper explores how the equity framework was deployed to better learn about a specific customer segment: customers that speak languages other than English (LOTE).¹ These are customers that may have based barriers to program participation due to the lack of resources in their preferred language. The first section highlights key aspects of the framework that are relevant to the examples detailed in this paper. What follows are three instances where the framework was applied, starting with the evaluation regarding the operationalization of efforts to improve language access services for customers. This will be followed by a discussion of how survey instruments, a common tool used by EM&V to gather information, were utilized to gather information on demand responses from non-participants. Another example of using a survey to better learn about LOTE customers will follow, yet in this case the addition of a language element may have caused some challenges, despite being successful in collecting information from these customers. The conclusion describes topics for future research based on these three evaluation activities, as well as initial observations on the deployment of the EM&V Equity Framework.

EM&V Equity Framework

The equity framework was developed by the EMC to provide guidance on (1) how equity can be addressed and approached in different stages of the evaluation process, and (2) how data collection efforts can be more equitable. The EMC used the framework to detail how equity can be engrained into each phase of the study process. Starting with the planning stage, once a study idea is approved for further research, the study working group will need to discuss the degree to which an equity lens can and should be applied to the study’s objective. From there, studies will fall into one of three categories, as described in Table 1 below.

¹ For this paper, the focus will be on LOTE customers that speak, read, write, or preferred to have services provided in Spanish, Portuguese, Haitian Creole, Mandarin, or Cantonese.

Table 1. EM&V Equity Framework Study Categories

Category	Description
Category 1	Studies that evaluate the PA's efforts specifically aimed at engaging customers and communities that have historically participated in PA programs at lower levels.
Category 2	Captures a broad range of work but is intended to ensure equity is integrated into the study process to a certain degree
Category 3	Studies primarily focused on producing factors that impact savings for PA programs

The equity framework notes that the integration of equity not only needs to be addressed when posing the research questions, but also when it comes to performing the work. While data and information gathering instruments will need to be enhanced to answer equity-related questions, in some cases the approach to engaging certain customers and populations needs to be taken into consideration. Therefore, the framework puts a particular emphasis detailing considerations around surveys, focus groups, and in-depth interviews since those methods allow for direct interaction with populations.

Surveys have been a common practice for EM&V work when attempting to better understand customer and community populations.² For studies addressing equity, whether directly or to some degree, the working groups for those studies that will conduct surveys will need to include questions to help answer equity-related topics. The working groups should also consider translating the surveys into at least one language commonly spoken in Massachusetts outside of English.

Just as it is important to develop survey instruments that contain equity-related questions, for surveys to be effective they need to be completed by customers and businesses of interest. Therefore, the equity framework encourages study working groups to recognize that reaching and understanding underserved customers and communities may require enhanced and community-centric approaches. Survey designs should integrate approaches such as multi-mode research and multi-lingual outreach to engage customers as needed. Researchers' multilingual and cultural expertise should be utilized along with local community organizations that already have trusted relationships with groups of interests. The framework notes that this is particularly significant for non-participation studies where the issues that have kept these customers from participating in programs could also prevent them from responding to surveys.

When it comes to applying the equity framework to learn more about LOTE customers, most studies will likely fall under Category 2. Yet there is a Category 1 study since the PAs are devoting resources to enhance language access services. The following section dives into this effort and how a new evaluation approach was utilized.

² Two of the examples described in this paper are around survey efforts, which is why equity considerations around this data collection method are described and the other two are not.

Language Access Plan

The PA's 2025-2027 plan put a specific emphasis on improving language access throughout the LOTE customer journey by utilizing trans-created³ material, interpreter services, and multilingual staff. These services are implemented under residential and small business programs under the auspices of the Language Access Plan ("LAP"). The LAP was put in place with a series of recommendations to help operationalize this effort, one of which was the need for ongoing collection of LOTE participant feedback.

The timing of the LAP development coincided with an ask made to EMC by policy stakeholders to focus on solutions-oriented evaluation approaches that could provide implementation with fast feedback. The result was an agreement between the EMC and PA staff involved with the implementation of the LAP to utilize an embedded evaluation. This evaluation approach allows evaluation activities to be integrated into the program implementation process as a way to establish a continuous feedback loop. It enables evaluators to have more direct access with implementors to collaborate and address challenges and topics in real-time rather than retrospectively.

The embedded evaluation process for the LAP primarily started as a consultative engagement as the plan was being operationalized. This consisted of evaluators participating in LAP implementation working group meetings and providing ad hoc commentary on topics they had experience with from other jurisdictions. The contribution and insights provided by evaluators through these working groups helped establish a level of collaboration and trust that would enable an embedded evaluation to be successful in the long run. This would become critical as the evaluation for the LAP moved into a stage where deliverables needed to be developed to help implementation address short-term challenges.

The first deliverable that evaluators helped LAP implementors with was by creating customer journey maps for the different programs that have a language access component. This originated when the different program implementors created a list of program documents that needed to be trans-created into different languages. The evaluators in the working group had already developed customer journey maps based on research they conducted with LOTE customers in Massachusetts. For the LAP, they leveraged that experience to come up with a matrix to help identify new gaps in documentation. In the spirit of the embedded evaluation's emphasis on timing and speed, the deliverable was created primarily by reviewing the list of documents provided by implementors.⁴ An illustration of this matrix is shown in Table 1 below.

³ Trans-creation is when a document, for example, in English is then taken and modified into a different language by not only translating it but also adjusting it to account for community and cultural significance.

⁴ Because embedded evaluations were established to provide implementors with real-time feedback, the EMC agreed that not all deliverables are developed for external reporting. This is why Table 1 is an illustrative version of the customer journey map and only notes certain sections of the document.

Table 2. LAP Customer Journey Map for LAP Trans-Created Documents (Illustrative)

Document Category	Program Awareness	Energy Assessment	Recommendations
Currently Exist	Website in English, Portuguese, and Spanish	Bilingual Staff	Translated templates in Spanish and Portugues
Gaps Previously Identified	Standardized translated communication materials	Track language services offered	Standardized resources
Wishlist Items	Translated program materials and leave-behinds	Translated FAQ documents	Local dialect translation support
New Gaps Identified	Translated marketing materials	Translated assessment checklists	Translated assessment reports and proposals

The customer journey maps created were successful not only in identifying gaps, but also getting LAP implementation buy-in and input. For each program, managers took the customer journey maps to their respective program teams to review and provide commentary and feedback. While this move might appear routine, the fact that implementation teams took the time to review it separately shows that (1) the level of detail provided by evaluators to implementors was appropriate, and (2) implementors saw the value of the work produced by the embedded evaluation. The latter point is important not only for the success of embedded evaluations, but also for the equity framework. Collaboration with implementation is an important consideration when reaching LOTE and non-participating customers for both program service delivery and evaluation purposes. The embedded evaluation approach can be a mechanism for that collaboration between implementation and evaluation that was highlighted in the equity framework to foster. At the time of this writing, the LAP embedded evaluation will be exploring a topic around survey collection that was brought about by the study described in the next section.

ConnectedSolutions

ConnectedSolutions is the active demand response (“ADR”) offering administered by four of the Massachusetts PAs (Cape Light Compact, Eversource, National Grid, and Unitil) for both residential and commercial customers. For the purposes of this paper, the focus will be on an evaluation of the residential program, where customers can participate using an eligible Wi-Fi enabled thermostat.⁵ In 2024, the EMC and PA ADR program implementors discussed conducting research around how to best incorporate ADR into customer segments where

⁵ Customers can also participate with behind-the-meter energy storage. However, this technology was out-of-scope for the evaluation described in this section.

participation has been a challenge. The idea was to better understand how to overcome barriers that keep these potential participants from enrolling in the program.

The EMC commissioned a study aimed at exploring how to increase participation in the residential ConnectedSolutions program through the use of wi-fi thermostats. The study consisted of program participation data analysis to see what populations were currently being served by the program, a review of existing enrollment pathways, and a survey of non-participants to understand their awareness of ADR and their willingness to enroll in the program. For the non-participant survey, the study team followed the guidance from the equity framework to decide that they would field Spanish and Portuguese version of the survey. Notably, the team also decided against sending out recruitment postcards in an effort to control study costs.

The non-participant survey was fielded to residential electric customers who have not participated in the ConnectedSolutions program. This included customers who may have received a rebate for a Wi-Fi thermostat through an energy efficiency program but did not enroll in ADR. Customers that completed the survey received a \$20 e-gift card. Survey questions were asked to learn about the type of cooling systems a customer has, as well as if they have a thermostat and if so then what type of device. The survey findings were presented to the EMC study team members, but the full anonymized data was provided to Eversource evaluation team members of only their customer responses.

The Eversource evaluation team analyzed the findings from 946 respondents to identify insights that could be informative to their ADR implementation team. Of those 946 customers, thirty-seven responded to the Spanish-language survey and just two responded to the Portuguese survey. Naturally, it would be difficult to draw any conclusions from the Portuguese survey. Yet from the Spanish-language surveys, over 70 percent of the respondents noted that they use a room, window, or through-the-wall air conditioner. From an implementation perspective, the findings from the Spanish survey would suggest that it would be difficult to get this specific set of customers to enroll into ConnectedSolutions because they would need to first have the central air conditioner or heat pump in their home. However, the English version of the study included a question about language preference that could lead to the opposite conclusion.

Figure 1 below shows the results of the type of cooling system owned by customers that responded to the Spanish-language survey versus those that completed the English version but noted that Spanish was spoken at home. As noted earlier, most of the respondents that completed the Spanish-language survey noted that their cooling system was a room, window, or through-the-wall AC unit. Yet based on the thirty-two⁶ responses from the English-language survey, one could be drawn to the exact opposite conclusion where the bilingual customer may be closer to becoming a ConnectedSolutions participant. However, the customers that completed the English version also identified other constraints and barriers that would prevent them from enrolling. Reasons highlighted in their responses include that they cannot get a Wi-Fi thermostat because they are renting their home, or that upgrading to a Wi-Fi thermostat is too expensive.

⁶ These thirty-two customers are those that completed the English-language survey and noted that Spanish is the only other language spoken in their household.

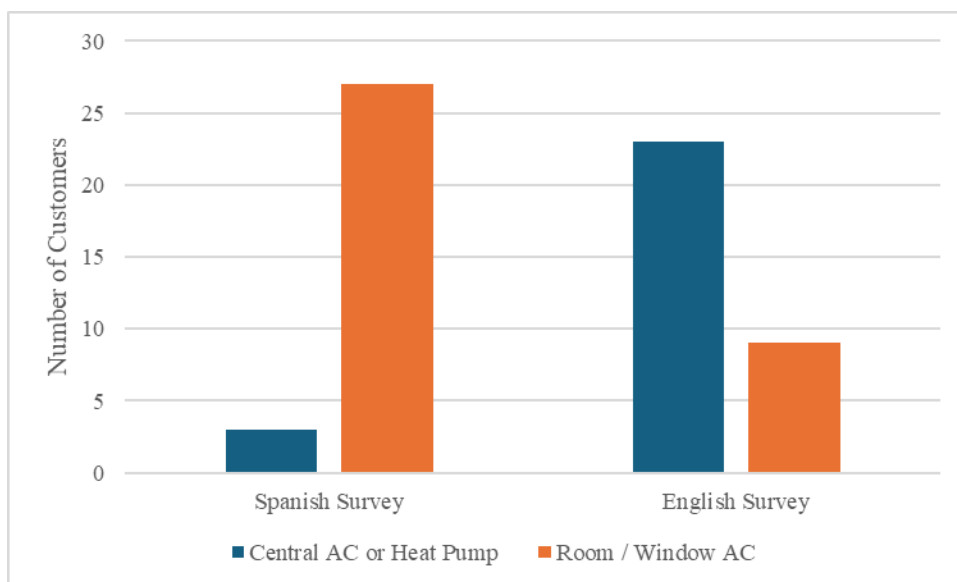


Figure 1. Type of Cooling System Indicated by Spanish LOTE Customers in Spanish- vs English-Language Survey.

Source: Guidehouse 2026

The results of the two surveys suggest more research is needed to better understand what is causing these two differences. While factors such as education and income could be one possible source, the Spanish-language survey did not gather enough information on either one to draw any kind of conclusion. Yet one outcome from this effort is the idea that Spanish-speakers do seem responsive to provide PAs with information through surveys. This will be a topic that will be explored through the LAP embedded evaluation, drawing from the experience with the ConnectedSolutions non-participant survey as well as the study described in the next section.

Residential Building Use and Equipment Characterization Study

Starting in 2017, the EMC has overseen research aimed at collecting saturation, characterization, and usage behavior data for electric- and gas-powered appliances, mechanical systems, and electronics in Massachusetts homes. This study, the Residential Building Use and Equipment Characterization Study (“RBUECS”), gathers this information through customer surveys and on-site metering. Work under the auspices of RBUECS has been done in phases, with the latest phase (“Phase 8”) taking place between March of 2024 and September of 2025.

Phase 8 of the RBUECS work introduced two modifications related to language. First was the development of a Spanish- and Portuguese-language version of the online survey. The second was refining the survey instrument to collect information on different demographic groups, including non-English speakers. The online survey was fielded focusing on two sets of customer groups. One was for customers that completed the 2023 RBUECS study survey with the goal of updating saturation data that was previously collected. The other was for new customers taking the survey for the first time. The latter group is where the language considerations were in effect.

Efforts to learn more about LOTE customers through this study did encounter challenges during the data collection process. From the translated surveys, twenty-eight responses were in Spanish and five in Portuguese, representing a small percentage out of the total 4,831 usable

responses. In addition, the introduction of language preference questions coincided with a decrease in the survey completed. The new customer survey was conducted in different waves, with wave four introducing with a language-preference component. However, approximately fifty percent of the respondents who completed that preference portion then went on to complete the full new customer survey. To compensate, the study team conducted additional waves of follow-up and expanded email recruitment to increase the total amount of responses. In total, the study team conducted eleven waves and contacted 225,000 customers. Table 3 shows the impact the language-preference component had on the survey completes.

Table 3. RBUECS Phase 8 New Customer Survey Recruitment

Wave	Customers Contacted	Survey Complete (Number)	Survey Complete (Percentage)
Pilot	5,000	103	2.1%
Wave 1	35,000	575	1.6%
Wave 2	35,000	575	1.6%
Wave 3	25,000	398	1.6%
Wave 4	25,000	218	0.9%
Wave 5	25,000	214	0.9%
Wave 6	25,000	294	1.2%
Waves 7-11	50,000	372	0.7%
Total	225,000	2,749	1.2%

Source: Guidehouse 2026.

Nevertheless, the intent of applying an equity lens was to provide some insight on LOTE customers for the first time in the history of this phased work. The RBUECS Phase 8 report produced weighted and unweighted survey demographics compared to the Massachusetts population for different categories, including language. Yet the raw data collected by the study provides further information around non-English-speaking customers. Collectively, information was collected on 4,831 customers, of which 658 noted that English is not their primary language. The analysis that follows will focus only on customers that speak just one of the languages prioritized in the LAP.⁷

The survey collected information from respondents on the type of building that they live in and whether they own or rent their home. Building type is classified by the number of residential units: one is single-family attached, two-to-four units is single-family attached, and five or more is multifamily. Figure 2 shows these results for customers that indicated that English is not their primary language as well as spoke one of the five languages prioritized by the LAP. For comparison, the same is displayed for customers that indicated English is their primary language. The raw data shows that survey respondents that spoke English, Mandarin, Cantonese, or Portuguese were more likely to own their home. On the other hand, Spanish speakers were more likely to rent. This highlights the possibility that Spanish speakers not only may experience

⁷ The only exception here is Haitian Creole, where two respondents indicated this was their primary language. Also, the survey the responses also picked up customers that are bi- or multi-lingual, or instances where just one language was spoken by a customer (e.g., Dutch, Farsi). Also, some Mandarin and Cantonese customers did not indicate which of the two options they preferred. Therefore, all Mandarin and Cantonese speakers have been combined into one category.

barriers to program participation related to language, but also as renters. This latter group may lack the decision-making ability to make energy efficiency improvements to their home.

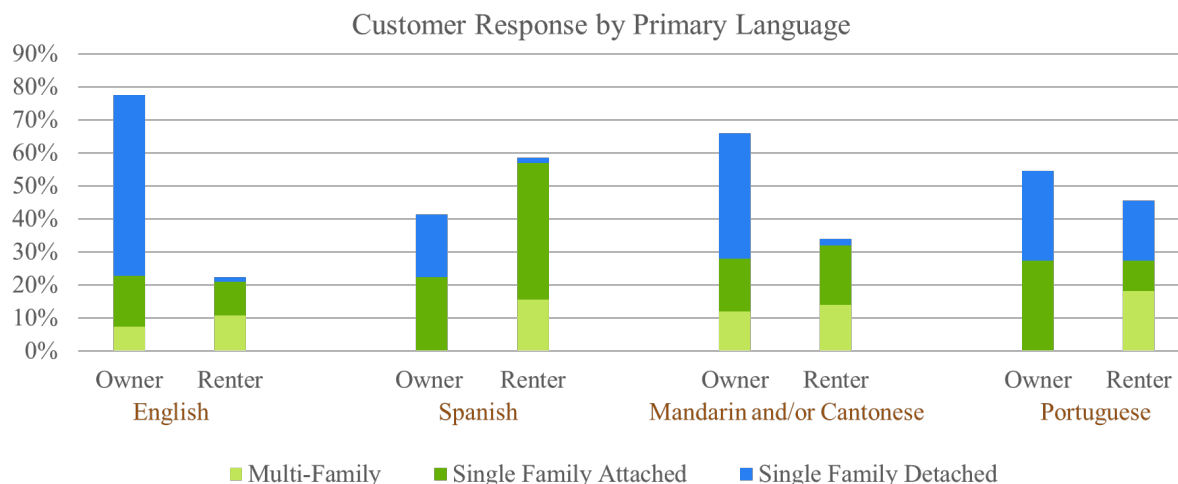


Figure 2. Type of residency and occupancy status of customers by language preference. *Source:* Guidehouse 2026.

The raw data collected also provided information as to whether a customer had previously participated in a PA-supported program. This insight can be helpful for PA implementors in better understanding their remaining market potential. Table 4 below shows the number of LOTE customers that noted that English is not their primary language, along with the percentage of those that participated in a PA program. The table also includes responses from those who indicated English is their primary language, as well as those whose answers indicate a degree of bilingual ability between English and Spanish.

Notably, the number of customers that participated in programs lags non-participants across all the language categories in Table 4. Yet customers that identified that Spanish, Mandarin, Cantonese, or Portuguese was their primary language had similar participation rates. Even with participation rates close to twenty percent, the analysis suggests LOTE customers are still behind those who primarily speak English. This would indicate that more efforts need to be made to increase not only LOTE customer participation, but also awareness of PA programs.

Interestingly, customers that noted some bilingual ability between English and Spanish had higher program participation rates compared to those who primarily spoke Spanish. The raw data suggests that these bilingual survey respondents were more likely to own their home.

Table 4. Language of LOTE Customer Program Participants and Non-Participants

Language	Number of Respondents	Percentage of Participants	Percentage of Non-Participants
English	3,875	32%	67%
Spanish	58	19%	81%
Bilingual – English and Spanish	69	39%	61%
Mandarin and/or Cantonese	50	18%	82%
Portuguese	11	18%	82%

Source: Guidehouse 2026.

Conclusions

The current term has provided the EMC with preliminary insights into LOTE customers as well as a clearer understanding of the evaluation practices used to gather information on this segment of customers. These efforts have also presented the committee with additional research opportunities. Both the survey instruments deployed for the ConnectedSolutions and RBUECS evaluations were successful in gathering some degree of information, particularly from the Spanish-language instruments. Conversely, the Portuguese surveys yielded little responses. As noted earlier in this paper, it could suggest that Spanish-speakers are more responsive to surveys in their native language, while the low participation from Portuguese speakers highlights a need for alternative engagement strategies. This topic will be evaluated through upcoming research aimed at better understanding what data collection methods work best for certain LOTE customers.

This follow-up research can hopefully address the possibility that a survey instrument could be a source of bias in and of itself. That is, if an evaluation study were to gather information from customers solely using surveys as opposed to canvassing or collaborating with community groups, then the data collection could be successful in trying to understand one specific group of LOTE customers at the cost of others. The outcomes of this research could help inform the EMC on how to design studies aimed at researching specific customer segments. Fortunately for the EMC and the PAs, the LAP embedded evaluation has established itself as a mechanism where language access questions can be explored in real time. This evaluation approach has been successful because implementation and evaluation came together over a mutual interest in better understanding LOTE customers. This new structure positions the EMC and the PAs to refine data collection methods and strengthen engagement with LOTE customers, creating a pathway towards more inclusive, equitable, and accurate program evaluation in Massachusetts.

The efforts from the ConnectedSolutions and RBUECS surveys to collect information on LOTE customers provided some initial observations on the deployment of EM&V Equity Framework. One is that the application of an equity lens could create time and resource constraints on a study. As noted earlier, the ConnectedSolutions study team decided against trans-created postcards to control costs. It's unknown whether these postcards would have

increased responses from Portuguese customers, but attempting to answer this as part of the ConnectedSolutions study would have been out-of-scope for that research.

Second, both surveys highlight how the EMC could be gathering a great amount of information on LOTE customers. Yet there needs to be a platform that allows data on LOTE customers to be analyzed. The RBUECS survey collected a considerable number of details on LOTE customers that was used by the report to determine saturation levels. The study report does did include an appendix where the raw data on LOTE customers could be found. Yet an analysis of this data to provide insights for PA implementors would have to be done outside of the RBUECS study.

The EMC's efforts to learn more about language access suggests that the LAP embedded evaluation serves not only to better understand how to best reach LOTE customers, but can also serve as a platform to analyze related information collected from other EM&V studies. This suggests that for an equity lens to be successfully applied to EM&V studies, it may require having mechanisms such as embedded evaluations specifically devoted to a specific population or topic of interest. That way, work on overcoming obstacles and barriers can be conducted within an appropriate scope and with dedicated resources. The approaches that come out of resulting research can then be applied within the EM&V portfolio where appropriate.

References

- Guidehouse. 2023. Massachusetts Limited English Proficient and English-Isolated Customer Journey Mapping and Barriers Study. Report prepared for the Massachusetts Electric and Gas Program Administrators. https://ma-eeac.org/wp-content/uploads/MA21R37-B-LEPJM-Report_FINAL_25Oct2023.pdf
- Guidehouse. 2026. Massachusetts Residential Building Use and Equipment Characterization Study Phase 8. Report prepared for the Massachusetts Electric and Gas Program Administrators. <https://ma-eeac.org/wp-content/uploads/MA24R69-B-RBU-Residential-Building-Use-and-Equipment-Characterization-Study-Comprehensive-Report-Ph8-Final-2026-02-03.pdf>
- Guidehouse. 2026. Residential Wi-Fi Thermostat Demand Response Participation Study. Report prepared for the Massachusetts Electric Program Administrators. <https://ma-eeac.org/wp-content/uploads/MA-Residential-Wi-Fi-Thermostat-DR-Participation-Study-Final-Report-2026-03-03-CLEAN.pdf>
- Mass Save. 2024. "The Massachusetts 2025-2027 Energy Efficiency and Decarbonization Plan" October 31. <https://ma-eeac.org/wp-content/uploads/Exhibit-1-2025-2027-Three-Year-Plan.pdf>