The Price Is Right: A Multi-Modal Approach to Researching Incremental Measure Costs

Ari Stern, Patrick Hewlett, Sam Sorrin, DNV, Medford, MA
Mike Rufo, Oakland, CA

ABSTRACT

As achieving cost-effectiveness becomes increasingly challenging in energy efficiency programs, the need for accurate incremental measure cost (IMC) data is growing. Such data is essential for comparing costs with societal benefits, including energy and demand savings, as well as non-energy benefits (NEBs). The data can support program efforts to set appropriate incentive levels and establish realistic program goals. However, not all data is reliable. Secondary cost data, for example, is often outdated and recirculated among sources without verification, especially for rapidly evolving technologies.

This paper assesses the strengths and limitations of primary research methods used to establish defensible IMCs. The research sought to achieve two primary objectives: 1) collect and analyze granular data on equipment costs and specifications to quantify the effect of energy efficiency on cost, and 2) collect localized cost estimates for a specific territory.

To develop the analysis, three primary data sources were used: national retailer websites, distributor price lists, and utility tracking data supplemented with AI-assisted lookups. We used a regression-based analysis enabled us to control for non-EE effects and isolate the specific cost increases associated solely with energy efficiency.

To capture regional cost variations, we conducted in-depth interviews with contractors exploring labor costs, ancillary materials, subcontractor fees, and markups. They also provided expert, qualitative insights into cost estimates.

Our findings help program administrators, policymakers, and evaluators invest in future cost research in the most economically and defensible manner. This data will provide more confidence in measure screening to ensure that program offerings are cost-effective to ratepayers and customers.

Introduction

IMCs quantify the additional cost to implement an energy-efficient measure relative to a baseline scenario. For example, installing a high-efficiency boiler is more expensive than installing a standard efficiency boiler. The IMC for a high-efficiency boiler reflects the additional equipment and labor costs beyond what a customer would have paid if installing a standard efficiency boiler.

A variety of programmatic planning purposes can benefit from good IMC data, including cost-effectiveness analysis and associated regulatory filings, integration with savings estimates (e.g., through a Technical Reference Manual), establishing incentive levels, estimating economic potential and setting associated goals, and integrated resource planning and GH mitigation planning. Given their importance in program planning, IMCs must be accurate and up-to-date. For example, in the scenario of LEDs becoming the standard practice, achieving portfolio-level cost-effectiveness in the energy efficiency and decarbonization landscape becomes more challenging. Accurate IMCs enable programs to right-size incentives and ensure that costs and benefits are accurately weighted in cost-effectiveness models.

In 2024, the New Jersey Board of Public Utilities (NJ BPU) and its Statewide Evaluator (SWE) team contracted with DNV to conduct primary research on IMCs for a prioritized set of measures (*NJ BPU 2025*). The study aims to update the IMCs used in the New Jersey Cost Test with New Jersey-specific data.

Additionally, new policy changes unlocked a suite of fuel switching measures in the portfolio; however, these measures lacked adequate IMCs that accurately reflected the actual costs of electrification. The prioritized measure groups included residential HVAC, residential water heating, commercial HVAC, and commercial refrigeration equipment. This work represents "Phase 2" of an overarching IMC effort for which Phase 1 leveraged secondary research to revise IMCs for all measures covered in the NJ TRM (*NJ BPU 2023a*). The secondary research effort found that measure costs are embedded in a variety of publicly available energy efficiency resources:

- **Potential studies:** Potential studies are typically conducted as part of a regulatory goal-setting process for energy efficiency portfolios. To estimate the potentially achievable savings from energy efficiency programs, the studies incorporate the incremental costs of various measures, along with other market forecasts. This often requires the creation of an incremental measure cost database (see next bullet).
- **Incremental measure cost databases:** IMC databases compare the costs of energy-efficient technologies to the costs of baseline alternatives, which vary depending on the measure implementation type (e.g., early replacement or replace-on-burnout). IMC databases broadly assign measure costs for entire portfolios of measures; in some cases, they also provide the sources used to estimate IMC for each measure. Occasionally, the sources used in the secondary research may be outdated, unavailable, or not transferable across jurisdictions.
- **Technical reference manuals:** TRMs provides standardized methodology to support how program implementers should estimate savings for various energy efficiency measures. Occasionally, TRMs include deemed IMC values for use in cost-benefit calculations. Like IMC databases, TRMs are comprehensive and can provide IMC for an entire portfolio of measures, based on the best available information at the time of publication. However, as with IMC databases, IMC values in TRMs often rely on secondary sources of varying vintage and availability. This can lead to a cycle of citations, where different documents mention the same sources year after year, resulting in values that appear recent but are based on outdated or unavailable primary data.
- **Evaluation reports and IMC studies:** Evaluation reports and IMC studies use primary research to estimate IMCs for specific measures. Reports are often published as directed by utility program evaluation plans and tend to focus on a narrow set of measures. IMC studies typically include secondary research, program tracking data and invoices, web-scraping data, and/or quotes solicited from market actors.
- **Benefit-cost ratio models:** Publicly available benefit-cost ratio (BCR) models, often in the form of Excel workbooks, show the results of benefit-cost tests used in energy efficiency programs and portfolio planning. These models often include measure-level costs but do not clearly cite sources for cost information.
- **Papers:** Academic/conference papers or whitepapers that review IMC methodologies. These typically do not have useful cost information, but they can offer lessons for future primary research.

The applicability of any of the above resource types may be limited for any given program. As mentioned above, some resources may cite data sources a decade old or more. Sometimes, the vintage of the original data source is obscured by circular references or references to a source that merely cited an even older source. The sources may also be representative of a specific geography that is not transferable to the utility program's market due to varying equipment costs, labor costs, and other market conditions. Sources may not have the granularity of data required to properly adjust costs to a program's context (e.g., data may not be split by equipment and labor, allowing for separate inflation and market adjustments, or by system quantity included in the project).

Since the Phase 1 study relied solely on secondary data with limited applicability to current New Jersey programs, the Phase 2 effort intended to improve the quality and applicability of incremental measure cost data for New Jersey for selected measures through primary data collection from manufacturers, wholesalers, contractors, and at the retail level of the supply chain using multi-modal data techniques.

Primary Data Collection

Phase 2 data collection primarily involved interviews with relevant market actors, such as contractors and distributors; "scraping" of retail cost data from online wholesalers; review of utility program tracking data; and examination of distributor price lists. The table below illustrates how the four primary data sources inform the key components of incremental measure cost values: equipment and labor costs for the high-efficiency system eligible for program incentives (i.e., "full costs") as well as for the baseline alternative.

Table 1. Primary data sources matched with incremental cost components

Data source	Baseline		Efficient	
	Equipment cost	Labor cost	Equipment cost	Labor cost
Market actor interviews	X	X	X	
Web scraping	X		X	
Program tracking data*			X	X
Distributor price lists	X		X	

* Program tracking data often did not distinguish between equipment and labor costs but combined the two into a total cost value for the high-efficiency equipment.

Each of the data collection techniques is further described in the following sections.

Program Tracking Data

DNV sought to collect actual costs for the installation of high-efficiency equipment rebated by utility programs. The requested tracking data covered project and measure-level data for all prescriptive measures rebated through residential or commercial programs between January 2022 and December 2023. In addition to project costs, DNV requested other relevant tracked information, such as project zip code, building square footage, savings claims by fuel, utility incentives, installation date, make and model, quantity, capacity, and rated efficiency. Where available, DNV requested that cost data be distinguished between equipment and labor to provide the maximum level of granularity for analysis.

In parallel with the New Jersey-specific utility data requests, DNV collected relevant program tracking data from other jurisdictions. Energy efficiency programs do not typically publish detailed tracking datasets, but some statewide initiatives (such as TECH Clean California; *TECH Clean CA 2025*) publish anonymized tracking data, including costs, for more transparency on market trends.

The usability of program tracking data varies widely by program and is dependent on the utilities' contractual agreements with implementers regarding data tracking. The most useful program data for incremental measure cost analysis includes the following:

- System-level costs as opposed to costs at the project level, which might group multiple measures into a single cost variable. If system-level costs are not available, system-type level costs (e.g., costs for all systems of the same type) can suffice if there is an associated quantity or capacity variable.
- Separate costs for equipment and labor allow for the inflation of equipment and labor costs to be tracked separately.
- Equipment capacity to normalize costs by capacity.
- Equipment efficiency to bin equipment into efficiency tiers, allow for regression analysis, and provide greater insights into cost differentials between equipment.
- Installation date to allow for the comparability of dollars across datasets.
- Model numbers allow for the confirmation of full equipment specifications.

It is important to conduct a due diligence examination of program tracking data to assess its reliability. While program data may include all the variables listed above, a closer examination of the costs may reveal signs of assumed cost values being used. For example, if there are few unique cost values relative to the number of projects or equipment models, or if the costs show unexpected correlations with capacity, the program may be using deemed values rather than actual values.

In some cases, program tracking data lacked the necessary additional variables to classify records and normalize costs, but it did have model number data. DNV used artificial intelligence to populate an empty table with relevant variables based on the model number, thereby enhancing the quality of program tracking data. Such an effort can also be employed to correct tracking data that may have incorrect entries for given variables (e.g., listing existing equipment efficiencies in the variable for incentivized equipment efficiency).

Lastly, as shown in Table 1, program tracking data can typically only be used to track the costs of efficient equipment, as the data only includes equipment that was incentivized through the program.

Market Actor Interviews

Across the four measure categories, DNV interviewed 41 market actors, primarily installation contractors . Subject matter experts developed interview guides tailored to the measures of interest as well as to the type of market actor interviewed. Each interview script gathered upfront information about the market actor's role in the New Jersey market and neighboring states and the type(s) of equipment they specialize in. The scripts next explored a selection of scenarios involving different equipment, event types, or building characteristics that could influence costs. The interview questions aligned with the relevant configurations specific to each measure group and addressed the cost implications of various configuration characteristics (e.g., pre-existing fuel type, the presence of ductwork, and the distance between the indoor head and outdoor unit).

DNV developed a sample of installers for each measure using program data, secondary research with North American Industry Classification (NAICS) codes, and additional research leveraging trade organizations. Recruitment happened via phone and email. For participants, DNV leveraged an introduction from the program staff to build trust with respondents. To complete interviews with busy contractors, interviewers needed to be flexible. Experienced interviewers were also responsible for scheduling their own interviews, which allowed them to be ready to conduct the interview the moment they got the contact on the phone the first time. Interviewers were prepared to conduct interviews on off-hours and throughout the day. For example, interviewers completed some interviews over the phone with contractors as they were driving between jobs.

Market actor interviews are uniquely well-suited for providing cost data specific to a program's jurisdiction. Still, care must be taken to limit the interviewee pool to market actors who work in the targeted area. To collect data useful for IMCs estimates, it is important to be selective about the equipment types asked about during the interview. For each equipment type, the interview should solicit costs for baseline efficiency equipment and one or two tiers of high-efficiency equipment. Providing a prototype project that details the equipment's capacity and efficiency can streamline the interview. The specificity of prototypical installations is crucial for commercial HVAC measures, where there is a high level of variation between projects in the market and a variety of specializations among market actors.

For the commercial HVAC measure group, which comprises hundreds of permutations of measures at various capacities and efficiency levels, DNV provided an additional incentive for installers to conduct a deep dive into their distributor portals. The deep dive consisted of a 3-hour webinar during which the installer shared distributor portals on screen while the team worked through dozens of costing scenarios. This provided a richer set of cost data for commercial HVAC. While not employed in this study, researchers could consider a more formalized relationship with a knowledgeable market actor, perhaps through a longer-term subcontract, to allow for a more comprehensive set of quotes for various equipment types, especially those where pricing is not readily available online outside of vendor portals.

Web Scraping

Web scraping involves the development and execution of automated scripts, or the use of artificial intelligence, to gather publicly available information online. The team collected equipment cost data from selected online wholesalers to quantify the effects of various equipment specifications on costs among numerous equipment models of varying sizes and efficiencies. Depending on the measure type and program design, DNV worked with Staff to select retail or wholesale websites to inform the material costs.

Scraped data also included characteristics unrelated to energy consumption, such as warranty period and decibel level, that could affect price. By employing artificial intelligence, DNV was able to cost-effectively web scrape a high volume of equipment cost data points, along with their associated characteristics of interest, such as capacity, efficiency, fuel type, and configuration, with minimal effort. DNV created an empty table with the variables of interest, populating the first column with the information required by the artificial intelligence tool to identify the source of each record. The tool was then able to populate the table with full specifications for each record in a matter of minutes. DNV confirmed the accuracy of this approach by manually confirming a sample of records. As discussed in the Analysis Methods section below, scraped data was primarily analyzed via hedonic price modeling to identify and quantify the characteristics that influence equipment costs, whether related or unrelated to energy efficiency.

Challenges with this approach arise from variation in variable naming conventions within a single retailer. Careful cleaning should be conducted to aggregate content into single variables. Additionally, since web scraping only provides equipment costs, it may be advisable to mark up these costs to account for the installer's markup. Installer interviews can shed light on how much equipment is marked up in the region. For this study, web-scraped equipment costs were marked up 10%-20% based on markup estimates determined from installer interviews.

Note that, unlike program tracking data, the unique count of records from web scraping represents the unique models offered by the retailer, not the unique count of systems actually installed.

Distributor Price Lists

Each market actor interview included a request for contact information for the contractor's primary distributor(s). Interviewers followed up with the named distributors to obtain a more comprehensive tabulation of equipment costs by configuration, manufacturer, and model. Using distributor price lists, DNV successfully collected prices for 125 unique models from three distributors of residential HVAC and residential water heating equipment. Distributor price sheets enabled analysts to compare costs with relevant equipment specifications by make and model, including capacity, efficiency, and cold-climate capability. However, manufacturers and distributors are typically hesitant or contractually prohibited from providing price lists, and thus, this type of data collection has a low response rate. As done with web scraping, it may be advisable to mark up equipment costs from distributor price lists to account for contractor markup. Similarly, to web scraping, the unique count of records from distributor price lists represents the unique models offered, rather than the unique systems installed.

Additional Secondary Research

DNV explored additional secondary research to fill gaps not covered by primary research methods. Gaps could include labor costs for specific measures, or equipment costs for large commercial measures, which, due to their custom nature, may have limited pricing availability online and in program tracking data. Below are three examples of secondary data sources leveraged in Phase 2.

- **RS Means** (*2024*) includes equipment costs and labor estimates specific to the NJ region; however, its granularity and transparency are limited. For example, equipment costs are typically not segmented by efficiency level in RS Means, so additional data sources were required for that level of granularity.
- **TRMs or cost databases.** Jurisdictions such as Pennsylvania (*PA PUC 2021*) included IMC data within TRMs or as an accompanying database to the TRM. These sources were leveraged in Phase 1 and revisited in Phase 2 to identify updated costs where available.
- **Cost research studies.** DNV identified recent cost research studies related to heating electrification, such as the Northeast States for Coordinated Air Use Management's "Heat Pumps in the Northeast and Mid-Atlantics: Costs and Market Trends" (*NESCAUM 2024*) and the Energy Information Administration's "Updated Buildings Sector Appliance and Equipment Costs" (*EIA 2023)*.

Analysis Methods

DNV first conducted outlier analysis on all collected data. Outlier analysis involved calculating interquartile ranges (IQRs) among the collected datasets and defining boundaries based on these IQRs. Outliers were defined as values below (Q1 – 1.5*IQR) or above (Q3 + 1.5*IQR). Any record identified as an o utlier was then removed from the datasets before proceeding to the next analysis steps. DNV analysts employed the three primary analytic techniques described below.

- **Hedonic price modeling** – Technology costs can vary by several parameters, only a subset of which relate to energy efficiency. Hedonic regression analysis quantifies price differences among influential parameters, allowing for the isolation of those attributable strictly to energy efficiency. Hedonic price models require large sample sizes with comprehensive data; web scraping and distributor price lists were the primary sources of such data. Hedonic price models ultimately produce correlations of equipment costs as a function of relevant independent variables (s), such as capacity and efficiency, controlling for all other non-EE effects, such as warranty level.
- **Matched-pair analysis** – Certain measures benefit from direct comparison between a high-efficiency technology and its standard-efficiency counterpart, preferably from the same manufacturer. IDIs with market actors, web scraping, and distributor price lists were the three primary sources that produced matched-pair data.
- **Simple averages** – Technologies with few non-efficiency attributes that affect price may allow for a more simplistic analysis approach. Data collection that produced a limited number of equipment attributes, such as market actor IDIs or program tracking data, primarily relied on this analysis method.

Data Synthesis

Given the variety of data sources described above, the team often synthesized different datasets to produce the most defensible IMC recommendation. The synthesis of results involved applying different weights based on the representativeness and applicability of the cost data to New Jersey. DNV collaborated with SWE to develop a weighting rubric for each data source available for each measure and efficiency tier. The weighting rubric considered the following:

- If a data source was not considered for a particular measure, a weight of 0% was applied.
- Greater weight was given to data source types that were available and credible for both the baseline and efficient cases. Examples include market actor IDIs, web scraping, and distributor price sheets.
- Greater weight was given to source types that represented relatively large sample counts, with consideration of the fact that the sample size is defined differently among the data sources. For example, in program tracking data, the count represents the total quantity of equipment rebated by the program. On the other hand, for web scraping or distributor price sheets, the count represents unique models, any of which may be broadly adopted by multiple customers.

General Analysis Assumptions

Incremental measure cost results incorporated the following parameter assumptions across all studied measures.

- **Equipment cost inflation** – DNV referenced the U.S. Bureau of Labor Statistics (BLS) Consumer Price Index to modify all equipment cost data to reflect 2025 dollars (*US BLS 2025a*).
- Labor inflation – DNV referenced the BLS Employment Cost Index to modify all labor cost estimates to reflect 2025 dollars (*US BLS 2025b*).
- **Discount rate** – Incremental costs from Early Replacement (EREP) of working equipment correspond to a more complex formula that accounts for the net present value of the existing working equipment. The net present value calculation incorporated an assumed discount rate of 6%.
- **Equipment lifetime** – The team referenced the NJ TRM (*NJ BPU 2023b*) for effective useful life values among all studied measures. Remaining useful life (RUL) values, which affect EREP scenarios, are generally based on the default one-third of the expected useful life (EUL).

Key Findings

Equipment costs are highly dependent on capacity, with larger systems costing more. Unsurprisingly, a major driver of equipment costs is the system's capacity. Therefore, it is best to normalize IMCs by capacity in this research.

For most equipment types, labor costs do not vary by efficiency level. In-depth interviews with installers and secondary research indicated that equipment efficiency does not impact labor costs for central air conditioners, heat pumps, storage water heaters, heat pump water heaters, or tankless water heaters. Labor costs differ between baseline and efficient systems when the efficient systems involve technologies that require additional labor to install. For example, efficient boilers are condensing, allowing for additional heat recovery from the exhaust, while baseline boilers are atmospherically vented. Condensing boilers require extra labor to install the necessary equipment for safely draining the condensate.

Program tracking data records often could not be used for incremental measure cost analysis. The quality and comprehensiveness of program data varied between and within utilities, programs, and measures. DNV encountered program records with no costs, deemed cost values, costs representing entire projects that could not be split by equipment type, and a lack of relevant variables such as equipment capacity or efficiency.

Small commercial HVAC systems have similar costs to residential HVAC systems. The commercial HVAC sector includes a small-capacity tier of residential-sized equipment with capacities of less than 5.4 tons. Installer interviews revealed close similarities in costs between small commercial HVAC systems and residential HVAC systems. Installers indicated that costs do not vary for residential-duty equipment installations regardless of building type.

Pricing data for large commercial systems is not as freely available online or in program tracking data as it is for residential-duty systems. HVAC systems with capacities over 5.4 tons are more likely to flow from manufacturer to distributor to contractor and are thus less likely to have pricing available on retail websites. Additionally, larger systems are less likely to be prescriptively priced but rather require customized quotes. Our analysis of program data revealed that the largest commercial systems likely follow custom pathways, making it difficult to normalize IMCs prescriptively.

Example Incremental Measure Cost Findings

Table 2 shows a sample of results from the NJ IMC study (NJ BPU 2025). It shows IMCs for Time of Sale (i.e., normal replacement) fuel switching scenarios. The study also included early replacement scenarios for each record. The IMCs for early replacement scenarios follow the same logic as described below, but are higher than Time of Sale scenarios because the early replacement scenarios have a lower baseline cost than Time of Sale scenarios (i.e., the present value of replacing baseline equipment in the future as is done for an early replacement scenario, is less than the present value of replacing the baseline equipment today.)

Rows 1 and 2 show scenarios in which a heat pump water heater replaces a gas water heater of similar size. The equipment costs are normalized by gallons, but the labor is by unit. This approach reflects installers indicating during interviews that labor costs do not vary by tank size. Equipment costs were estimated separately for rows 1 and 2; however, it makes sense that the equipment cost per gallon was lower for row 2 because row 2 involves larger tanks, and there are cost efficiencies associated with scaling up.

Rows 3 and 4 show scenarios in which a ducted air source heat pump replaces a central air conditioner and furnace combination, with Row 4 corresponding to a higher-efficiency ASHP than Row 3. Both equipment and labor costs are normalized by tons since the installer indicated that both vary by system capacity. Row 4 has higher equipment costs than Row 3 because the higher efficiency is associated with higher equipment costs. The labor costs are the same between Rows 3 and 4 because the difference in efficiency does not impact labor. Still, there is an incremental labor cost for installing an air-source heat pump relative to a central air conditioner and furnace, as some air-source heat pump projects require an electric panel upgrade.

Rows 5 and 6 show scenarios in which mini split heat pumps fully replace heating and cooling provided by a boiler and room air conditioner(s). Again, there are two scenarios, with Row 6 representing a higher-efficiency mini split than Row 5. As with air-source heat pumps, the equipment IMC varies by efficiency, with Row 6 (the more efficient mini-split) having a higher equipment IMC than Row 5. The labor IMC does not vary by efficiency but is substantially higher than for Rows 4 and 5 because mini splits require more labor than the baseline counterpart due to the configuration of long refrigeration lines and other individual pieces of equipment.

Rows 7 and 8 show the same scenarios as Rows 5 and 6, except that Rows 7 and 8 represent partial replacements rather than full replacements. These scenarios are akin to a household installing a single mini split to add heating or cooling to a part of their home, such as a three-season porch, while keeping the primary heating system, a boiler, in place. The baseline case includes the cost of adding heat to the porch by extending the boiler's distribution system and installing a room air conditioner for cooling. We assumed that extending the boiler distribution system would incur negligible material costs but would require one-third of the labor cost of installing a new boiler. Additionally, the baseline case includes the present value of future boiler replacement needs. The estimated IMCs for rows 7 and 8 are higher than in rows 5 and 6 because the baseline costs are lower for rows 7 and 8 (it costs more to fully replace a boiler and room air conditioner than it does to extend a boiler's range and replace it).

Table 2. Primary data sources matched with incremental cost components (time of sale scenarios)

#	Fuel Switch Type	Baseline case	Efficient Case	Equipment IMC	Labor IMC
Heat pump water heaters					
1	Full	Gas storage water heater (UEF < 0.64, gallons <=55, atmospherically vented)	Heat pump water heater (gallons <=55, UEF>=3.3, 240V)	$33.35/gal	$596.25/unit
2	Full	Gas storage water heater (UEF < 0.78, gallons > 55, atmospherically vented)	Heat pump water heater (gallons > 55, UEF>=3.3, 240V)	$23.48/gal	$596.25/unit
Ducted air source heat pump					
3	Full	Central Air Conditioner (13.4 <= SEER2 < 15.3) + Furnace (AFUE < 95)	Air Source Heat Pump Tier 1 (15.3 SEER2 12.4 EER HSPF2 8.5)	$1,226.22/ton	$349.62/ton
4	Full		Air Source Heat Pump Tier 2 (17.2 SEER2 12.4 EER HSPF2 8.5)	$1,807.75/ton	$349.62/ton
Mini-split heat pump					
5	Full	Boiler and room air conditioners (80 AFUE & 8.6 EER2)	Mini Split Heat Pump Tier 1 (17.2 SEER2 11.7 EER2 8.5 HSPF2)	$1,987.70/ton	$1,291.13/ton
6	Full		Mini Split Heat Pump Tier 2 (19.3 SEER2 11.7 EER2 10.8 HSPF2)	$2,193.25/ton	$1,291.13/ton
7	Partial	One third of labor costs for boiler + full boiler equipment and labor costs for future replacement + room air conditioner costs (80 AFUE & 8.6 EER)	Mini Split Heat Pump Tier 1 (17.2 SEER2 11.7 EER2 8.5 HSPF2)	$2,095.21/ton	$1,301.38/ton
8	Partial		Mini Split Heat Pump Tier 2 (19.3 SEER2 11.7 EER2 10.8 HSPF2)	$2,300.76/ton	$1,301.38/ton

Lessons Learned and Conclusions

When planning a cost estimation study, it is essential to be selective about what the study aims to accomplish in terms of measures, permutations of those measures, and how the results should be used. With a limited budget, only a limited number of measures can be investigated in sufficient detail. The more permutations of measures chosen, the more difficult it will be to get a large enough sample of cost estimates from installers, and the less likely you are to find sufficient granular data from other sources. In regard to how the data will be used, if the study aims to inform incentive amounts for normal replacement measures, the study needs only to estimate IMCs. However, if the study will be used to inform early replacement measures, the study must then also get full costs for all incentivized and existing equipment types. Study designers should use programs planned goals and participation levels by measure to prioritize measures for research.

The following key lessons from this effort can inform future cost estimation studies:

- **Web scraping is the lowest-cost and most comprehensive method of data collection and works particularly well for residential measures.** We leveraged artificial intelligence to gather costs and full specification data for both baseline and efficient equipment. This comprehensive set of data enabled regression analysis to isolate the impact of efficiency on price. Web scraping is less valuable and less relevant for commercial purposes, as companies are less likely to disclose pricing publicly and are less likely to follow a traditional retail supply chain.
- **Program tracking data had limited usefulness.** Program data can only provide costs for efficient equipment. Moreover, program data often has deemed costs, costs that cannot be attributed to specific systems, or a lack of additional variables required to normalize costs or bin by efficiency. Adequate data is particularly sparse for commercial equipment, which is more likely to follow custom pathways than residential equipment. If model numbers are provided, artificial intelligence can be easily deployed to supplement program data or confirm the accuracy of program data.
- **Distributors and manufacturers are reluctant to provide distributor price lists.** Distributors' price lists are not publicly available, and distributors are often reticent or legally prohibited from providing such data. In our study, we were able to leverage personal connections to obtain a handful of distributor price lists; however, a study should not rely solely on distributor price lists to achieve full coverage of targeted measures.
- **Funneling a concentrated incentive to a small number of contractors is an effective way to gather data for commercial measures.** As discussed above, commercial cost data is not easily obtainable from web scraping, program tracking data, or distributor price lists. This is because commercial equipment, especially large equipment, is more likely to have custom pricing and follows supply chains that are less publicly transparent. We mitigated this problem by providing a larger incentive of $500 to a single installer to do a deep dive with us on commercial equipment. The installer shared their screen during a two-hour online interview. We asked the installer about specific equipment types, and they showed us the costs and associated specifications in the various distributor portals to which they had access.
- **Interviews with commercial installers should ask about specific prototype installations.** Commercial equipment can be highly customized and installers in the market can be highly specialized. To get a useful sample size of comparable costs across installers it is important to specify the type, size, and location of a prioritized measure and to asks for cost estimates for specified efficiency levels. Costs should be requested separately for equipment, associated materials, and labor.

References

EIA. 2023. "Updated Buildings Sector Appliance and Equipment Costs and Efficiencies."

NESCAUM. 2024. "Heat Pump in the Northeast and Mid-Atlantic: Costs and Market Trends."

NJ BPU. 2025. "X2218 New Jersey Energy-Efficiency Incremental Measure Cost (IMC) Study Report: Phase 2 – Primary Research."

NJ BPU. 2023a. "X2218 New Jersey Increment Measure Cost Study Phase 1."

NJ BPU. 2023b. "New Jersey 2023 Triennial Technical Reference Manual."

PA PUC. 2021. "2021 Technical Reference Manual."

RS Means. 2024. "Construction Cost Data." https://www.rsmeans.com/. Accessed May 2024.

TECH Clean CA. 2025. "Heat Pump Data." https://techcleanca.com/heat-pump-data/. Accessed August 2024 and June 2025.

US BLS. 2025a. "Consumer Price Index." https://www.bls.gov/cpi/. Accessed May 2025.

US BLS 2025b. "Employment Cost Index." https://www.bls.gov/eci/. Accessed May 2025.