Learning by Doing: Contractors' Learning in Heat Pump Installations *

Christopher R. Knittel,[1,4] Johnattan Ontiveros,[2] Juan Palacios,[2,3] and Siqi Zheng [2*]

[1] Sloan School of Management, Massachusetts Institute of Technology

[2] MIT Sustainable Urbanization Lab and Center for Real Estate, Massachusetts Institute of Technology

[3] School of Business and Economics, Maastricht University, the Netherlands

[4] National Bureau of Economic Research

November 30, 2023

Abstract

This paper examines the role of contractors in the adoption of high-efficiency heat pumps in the residential sector, a critical building decarbonization technology in the energy transition of households. We compile a unique dataset that includes comprehensive information on the characteristics of installed equipment as well as the cost breakdown in around 14,000 heat-pump installations, undertaken by 600 different installers between 2015 and 2019 in the state of Massachusetts (US). We document large systematic differences in installed capacity and costs across contractors, after controlling for a rich variety of home attributes. In addition, we document a significant learning effect in heat-pump installations, characterized by a reduction in the degree of oversizing installations as contractors accumulate more experience. An increase of 1% of the number of installations by a contractor is associated with a 0.03% reduction in installed capacity of heat pumps. This learning effect is more salient for the less complex installation tasks such as single-head heat pumps and those in the homes previously using natural gas heating systems.

*The authors gratefully acknowledge the support from the MIT Energy Initiative. We received very helpful comments from MIT-CEEPR Economics Seminar.

1 Introduction

The United States and the European Union are under mounting pressure to fulfill their commitments to reducing greenhouse gas emissions in the coming decades, including their ambitious net-zero targets by 2050 (Kerry, 2021; European Comission, 2020). Building decarbonization plays a key role in energy transition pathways towards a net-zero economy in both regions. The building sector represents approximately one-third of national emissions (International Energy Agency, 2022c), with on-site heating systems powered by fossil fuels (71% of current stock) representing one of the largest barriers to a net-zero society (International Energy Agency, 2022b).[1]

High-efficiency electric heat pumps, which provide the most energy-efficient solution to space heating, are considered the most promising solution to reduce emissions from the building sector (International Energy Agency, 2022b). According to the International Energy Agency (2022a), the number of heat pumps globally must increase from 180 million (2020) to 600 million installed units by 2030 to support the net-zero targets pledged by major economies, tripling the existing space heated by this technology. Currently, the adoption of heat pumps is hindered by the higher upfront costs of installation compared to gas and oil heating alternatives and the lack of qualified and experienced installers, who are in great need of skilled labor to meet the demands of this emerging and essential technology (International Energy Agency, 2022a). The presence of unqualified installations may introduce a risk of inefficiencies in the equipment and extra operational costs driven by inappropriately sized equipment.

In this study, we compile a unique data set containing records of all heat pump installations in Massachusetts over a five-year period (2015-2019) to explore the installation practices of contractors and how they conduct "learning by doing" as they gain experience in installations. Our data set includes detailed characteristics of the dwelling and the type of equipment installed, along with the identity of the contractor who completed the installation. In addition, for each installation, the data include the total cost of installation and the capacity of the heat pump installed, both of which are important indicators for the upfront cost and quality of installation. Installing a heat pump that is too small or too large for a given house results in losses in energy efficiency or comfort and increased operational costs.

We document several empirical results. First, we find that the installed heat-pump capacity and the costs of installations per property vary significantly from contractor to contractor after controlling for a rich set of property and heat-pump equipment characteristics. Second, our main analysis shows that contractors' experience plays a significant role in explaining these differences. We analyze the association between experience, measured by the number of past

[1] The current building portfolio includes 61% of buildings using natural gas and 10% using oil to fuel their properties.

installations a contractor has completed, and the cost and heat pump capacity of the installation while controlling for the extensive set of dwelling and heat pump characteristics. Then, using a fixed-effects model exploiting the within-contractor variation changes in costs and installed capacity; we show that as contractors accumulate experience, the sizing of heat pumps decreases significantly. An increase of 1% in the number of installations by a contractor is associated with 0.03% reduction in the installed capacity of heat pumps after controlling for a comprehensive list of housing and heat pump characteristics. This result suggests that contractors may first choose oversized heat pumps, and as they learn over time, they are downsizing the heat pumps they install to fit the heat pump into the house more accurately. The change in the size of heat pumps is the main determinant for the changes in installation costs.

Our paper builds on the existing literature concerning contractors in the installation of energy-efficient equipment for households. In their seminal paper, Allcott and Greenstone (2012) introduce the issue of the “Energy Efficiency Gap,” which describes the systematic discrepancy between the engineer predictions and realized performance of energy efficiency investment. The authors identify two types of market failure that bring about the gap: energy use externalities, coming as a result of adjustments in behavior from energy efficiency improvements completed in a home (“rebound effect”), and investment inefficiencies in the form of foregone profitable improvements arising from imperfect information or inattention (Allcott and Greenstone, 2012). In a recent paper, Christensen et al. (2021) estimate the sources of energy-efficiency performance gaps from projected and realized energy savings from the Illinois Home Weatherization Assistance Program. The authors found that 41% of the gap was due to bias in the engineering model and 43% due to heterogeneity in contractor workmanship throughout weatherization installations (Christensen et al., 2021).

In addition, it contributes to the literature estimating the learning curves of professionals. There is extensive literature on the effects of learning on the quality and productivity of workers, often referred to as 'learning by doing' (for a review, see Thompson, 2010). Specifically, these models refer to a form of passive learning acquired from completing successive tasks and 'improving' over iterations. Interestingly, research done by Jovanovic et al. (1995) structures learning in a Bayesian framework, where a superior technology exists, but the adoption first requires an accumulation of experience to overtake inferior technologies currently in use. This draws clear parallels to the transition from existing and familiar gas and oil-based heating to highly-efficient heat pumps that many contractors remain inexperienced in installing. Similar work has further been produced for the early introduction of solar photovoltaic installations in California, where the learning of contractors over a ten-year period was found to reduce costs of installation (Bollinger and Gillingham, 2019).

Our paper speaks to the above literature by examining the role of contractor characteristics and experience in explaining the energy efficiency gap. Past literature identifies the gap between expected and realized returns from an investment in energy efficiency, which poses as challenges to achieving widespread adoption of energy-efficient technologies. We identify an important role contractor experience plays for installation costs, but specific patterns in installation practices, along with the effect of learning with this continuously emerging technology, have yet to be examined in depth in the literature. Our paper utilizes a new data set of heat pump installation records that is unique in its amount of information and span of homes to answer the above questions.

The paper proceeds as follows. Section 2 describes the process of installing a heat pump and the landscape for installer certifications and training programs. Section 3 describes the data compiled for the analysis, followed by Section 4 that presents the empirical model, and Section 5 describes our empirical results. Section 6 includes the conclusions and directions for future work.

2 Required Skillset and Training for Heat-Pump Contractors

The deployment of a heat pump in residential properties is a demanding exercise that requires a highly skilled HVAC contractor to do a variety of tasks, some of which are novel to experienced contractors who have primarily worked with gas and oil-based heating in the past. First, the installation of a heat pump requires a more complex assessment of the right sizing than the evaluations commonly used for gas heaters, requiring that factors such as the quality of the building's insulation be accounted for rather than solely the square footage of conditioned space (Gleeson, 2016).

Additionally, the installation of heat pumps tends to require the upgrade of existing elements of the heating system and/or the electric system of the property. For this, installers need to assess the expected energy demand for the climate of the location. An improper assessment of the heating needs of the building can have a substantial impact on the performance of the heat pump (Domanski et al., 2014). In a recent study done by the U.S. Department of Commerce, the faults with the most potential for performance degradation and increased annual energy consumption were duct leakages, refrigerant under or overcharge, oversized heat pumps, and low indoor airflow due to undersized ductworks (Domanski et al., 2014). A proper installation, therefore, requires a nuanced understanding of a building's heating needs, its current infrastructure compatibility, and good judgment on whether further modifications, such as the improvement of insulation or resizing of existing ductwork, is needed.

The shortage of active training programs for contractors makes them rely on passive on-the-job strategies to learn how to select the right-sized heat pump for houses. There are limited certification requirements for HVAC contractors that install heat pumps. In Massachusetts, for instance, there are no required HVAC certifications for practice, with optional national certifications such as the North American Technician Excellence (NATE) certification (US EPA, 2015).[2] These certifications do not overlap with the certifications required for fossil fuel-based heating: a license issued by The Board of State Examiners of Plumbers and Gas Fitters to Journeyman or Master plumbers for gas heaters, or a Oil Burners Technician Certificate for oil heaters (Mass.gov, 2023b,a).

3 Data Description

We compile a unique data set that includes comprehensive information about 20,000 heat-pump installations in Massachusetts, including the characteristics of the heat-pump equipment installed, the house, and the contractor who conducted the installation.[3]

We obtained all heat pump installation records in the state of Massachusetts between 2015 and 2019 from the Northeast Energy Efficiency Partnership (NEEP). NEEP compiled installation records from all homeowner rebate applications. Each installation record contains the costs and relevant characteristics of the home and installation. In particular, the total cost of installation, the identity of the contracting company, the number of heat pump units used along with details on the equipment used (i.e., brand and model of the heat pump, and installed capacity), the living area, and retrofit status of the installation are recorded. The completed data set consists of 18,804 installations within Massachusetts from the years 2015 to 2019, executed by 627 unique contractors.

In addition, we collect a comprehensive list of hedonic characteristics of each house in our sample from web scraping Zillow, a leading online real estate marketplace that sources data on properties from a mix of public records and seller/owner data input. [4] For each property in the sample, we collect its size, type (e.g., single-family or multifamily), year of construction, number of bathrooms and bedrooms, value, and transaction records.

Finally, we collect data on each contractor company in our sample. This information is gathered for the contractors present in the heat pump data set through Data Axle, a data provider for US and International companies. The company data include factors such as the year of the company's founding and the size and revenue of the contracting companies.

[2]In order to work with refrigerants, which are used in the installation and maintenance of heat pumps and air conditioners, Massachusetts adheres to the EPA section 608 Refrigerant Certification which is received upon scoring a passing grade on the section 608 examination US EPA (2015).

[3]Cleaning and transformation of the data are specified in section 7.1.

[4]Zillow is distinguished as a data source by its breadth, containing information on over 135 million US homes Dun & Bradstreet (2023).

3.1 Summary Statistics

Table 1 provides the summary statistics of our final data set for the analysis. Panel A in Table 1 describes the characteristics of each installation in our sample. In total, we have information about 14,369 installations from 2015 to 2019 (after dropping data points with missing values). The average cost per installation is USD 9,367. The average installation has 1.3 heat-pump units and 29,157 BTU of installed capacity. The vast majority of these installations were retrofits to existing homes with a previous form of heating (natural gas heating for 43% and oil-based heating for 38%).

Panel B in Table 1 describes the characteristics of houses in the sample. The average size of a dwelling in our sample is 1,894 square feet, with 3.2 bedrooms and 2 bathrooms. Houses in our sample were built between 1714 and 2019, with the average house in our sample built in 1959. The virtual totality of heat pumps in our sample (99%) are installed in existing houses, with only 1% of heat pumps installed in newly built homes. The majority of houses in our sample are single-family houses (93%), with only 6% of heat pumps being installed in multifamily or condo units.

The average contractor company in our data set has installed 75 heat pumps, with a range of installations going from 1 to 606 heat pumps. The companies are around 40 years old on average, with large differences in year of establishment (from 1872 to 2021). Finally, the ranges from zero to 606 registered employees, with the average contractor company having 11 registered employees.

3.2 Graphical Evidence

Distribution installations. Panel a in Figure 1 presents the time series of installations in our sample. There is an increasing trend in installations over our sample period. There are 2,614 installations recorded in 2015 in our data and 6,258 installations in 2018. In addition, heat pump installations have a seasonal trend (Fig 1 [a]), peaking in the summer months and dipping in the winter months.

Panel b in Figure 1 displays the regional distribution of installations in our sample. There is a particularly high uptake of heat pumps in Barnstable, Middlesex, and Worcester counties by pure counts.

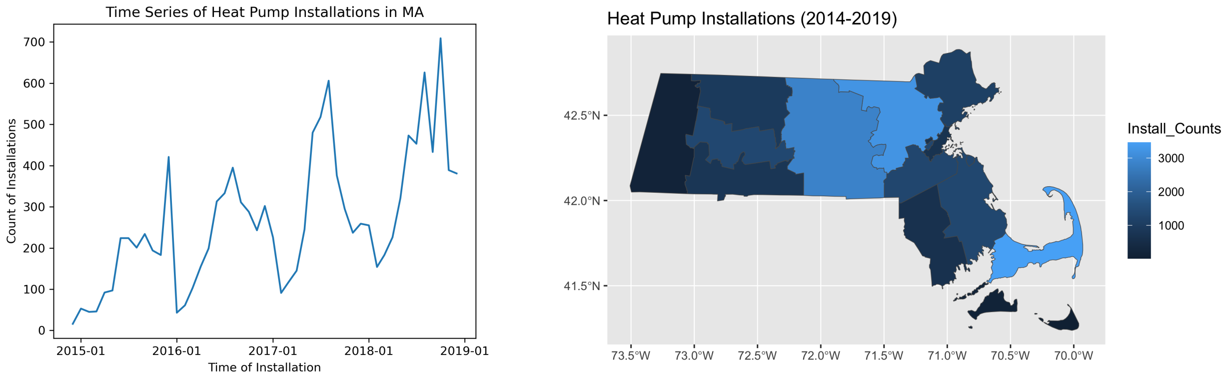
Distribution installation costs. Using a subset of MassSave data that includes the details of the breakdown of cost (N=163), figure 2 provides a visualization of the categories of cost for a heat pump installation in single-family homes. Besides the cost of the heat pump unit itself, the cost of the labor to install is the largest source of cost in an installation, potentially exacerbating the high upfront cost barrier to widespread heat pump adoption.

Table 1: Summary Statistics

Variables	***Count***	***Mean***	***Standard Dev.***	***Min***	***Max***
(A) Heat-Pump Installations					
Year of Installation	14369	2017	1.2	2014	2019
Total Cost of Installation (USD)	14369	9367	5449	1000	53000
Number of Heat Pump Units	14369	1.3	0.6	1	5
Installed Heating Capacity	14369	29157	13809	6968	144000
Original Fuel Type					
Natural Gas	5350	0.43	-	-	-
Oil	6052	0.38	-	-	-
Electric Resistance	1842	0.13	-	-	-
Propane	636	0.05	-	-	-
Wood Stove	246	0.02	-	-	-
(B) Home Data					
Living Area (Sqft)	14369	1894	1074	260	59567
Number of Bedrooms	14369	3.2	0.8	1	5
Number of Bathrooms	14369	2	0.8	1	5
Year Home Built	14302	1959	38	1650	2022
Last Selling Price	11420	325765	258638	1	6232500
Assessed Value of Property	14369	612746	367555	130300	9326400
Home Type	14369	-	-	-	-
Single-Family	13392	0.93	-	-	-
Condo	432	0.03	-	-	-
Multi-Family	422	0.03	-	-	-
Apartment	53	0.00	-	-	-
Unknown	14	0.00	-	-	-
Townhouse	31	0.00	-	-	-
Lot	10	0.00	-	-	-
Manufactured	15	0.00	-	-	-
(C) Contractor Data					
Year Established	147	1978	32	1872	2021
Number of Employees	305	11.6	14.8	0*	100
Number of Previous Installs.	644	15	30	1	296

*Owners of contracting companies are not included in employee counts.

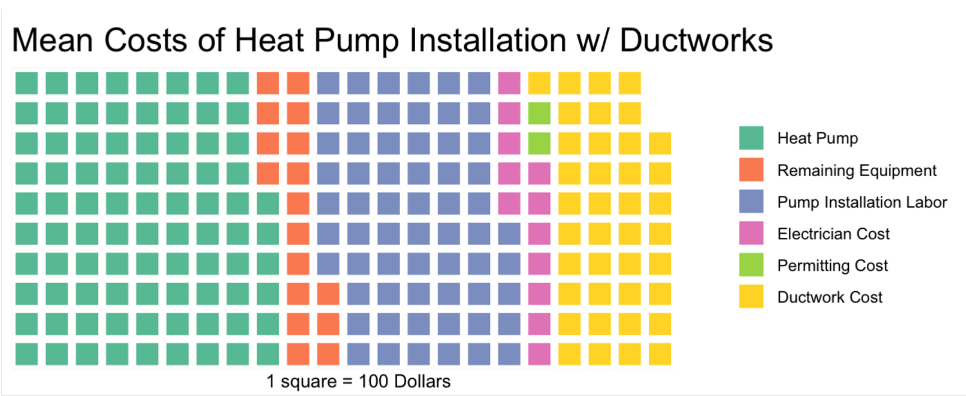
Figure 1: Count of Heat Pump Installations over time (a) and by Massachusetts County (b)



(a) Time series of installations

(b) Regional distribution installations

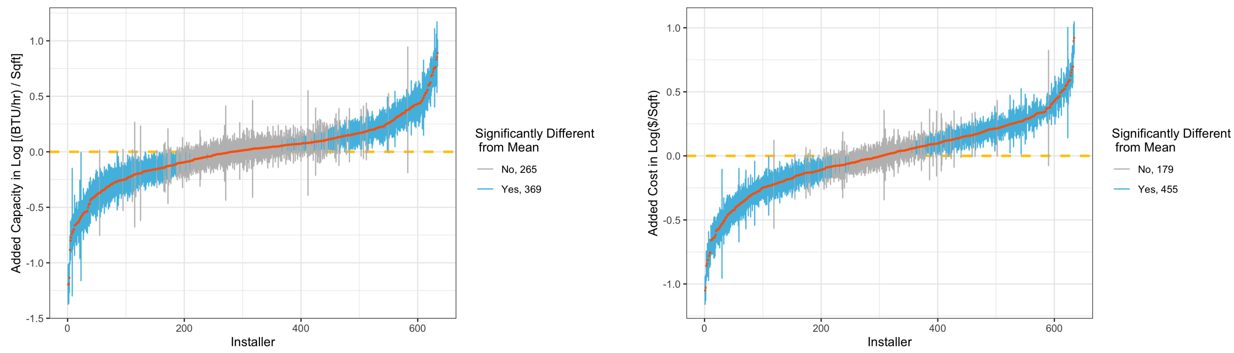
Figure 2: Cost breakdown for an average heat pump installation with ducting for a single-family home.



Next, we explore the systematic differences across contractors in pricing and installed capacity. Figure 3 depicts the fixed effects associated with individual contractors in a regression with all hedonic characteristics of the property where the heat pump is installed to control for idiosyncratic differences in the sample of houses where contractors operate. Each point represents the contractor-specific average estimated added cost, holding all other control variables constant. Bars represent the 95% confidence intervals of costs for each contractor. The results are additionally standardized around the mean across contractors. Contractors with intervals colored in blue indicate contractors who have an estimated added cost that is significantly above or below the mean. Lastly, the dependent variables of total cost and installed capacity are standardized by the square footage of the home and log-transformed.

For a contractor's added cost, the majority of contractors significantly deviate from the mean (455 compared to 179). For the 25th percentile, the added cost is a 14.28% decrease compared to the mean across all contractors. For the 75th percentile, the added cost is 20.32% increase from the mean. Equivalently, for an average house of 1,894 square feet, with a total cost of 9,367 USD, the differences between the 25th and 75th percentile of contractor effects are $\$4.2394/Sqft$ and $\$5.9505/Sqft$ or $8,029.42$ USD and $11,270.25$ USD. For the contractor's effect on installed capacity, 369 contractors are estimated to deviate significantly from the mean. The 25th percentile of estimates is a 13.70% decrease from the mean, and a a 16% increase for the 75th percentile.

Figure 3: FE Estimates by Contractor on Total Cost and Heating Capacity



(a) Added Heating Capacity

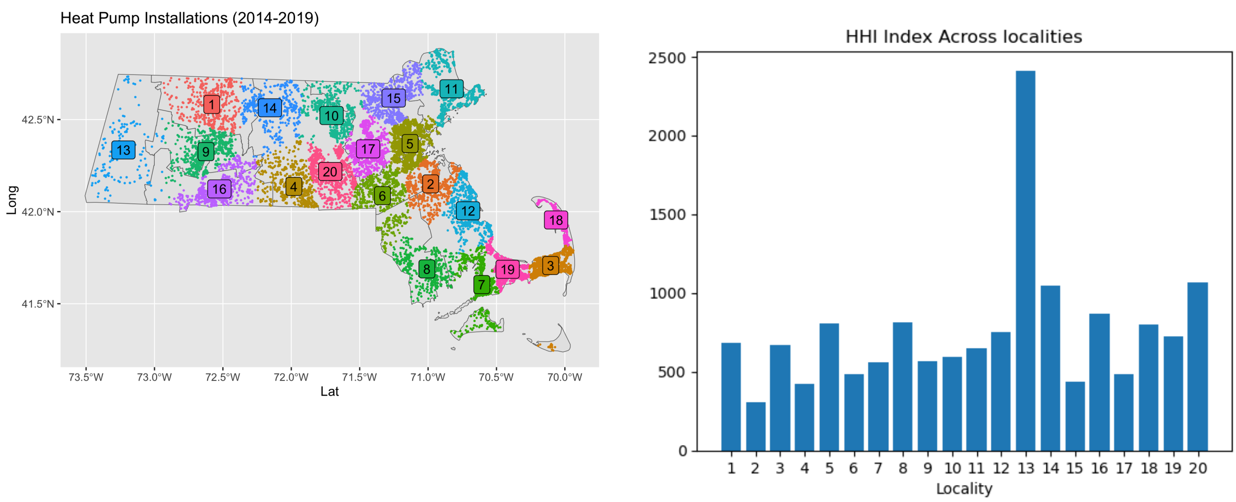
(b) Added Cost

†The Y-axis is on the log scale for (Capacity/Sqft) and (Cost/Sqft). The individual estimates x are interpreted on in un-logged terms by taking the exponent (e^x), to get the multiplicative effect on the dependent variable. An estimate of -0.1541, translates to $exp(-0.1541) = 0.8572$ multiplicative effect or equivalently 14.28%

3.3 Competition Within Local Markets

A final aspect of the data to be explored is the spatial distribution of heat pump installations within Massachusetts. The eyeball observation is that there are clear clusters of installers' contracts, and it seems that many of the installers operate quite locally. This bears interest in the context of contractors because contractors who operate near each other are presumably competing in the same market for installations. In order to measure this, we use the HHI index, a commonly used measure of market concentration (US Department of Justice, 2015). HHI values below 1,500 are considered unconcentrated industry, and HHI ranges between 1,500 and 2,500 are considered to have moderate concentration (US Department of Justice, 2015). Here, we group the heat pump installations into 20 sub-markets through K-means clustering specified in section 4, and calculate the HHI index for each sub-market, as shown in figure 4.

Figure 4: HHI Index Across selected 20 Sub-Markets



(a) Locality Clustering Map (b) HHI Values per Locality

Colors in figure (a) represent the localities in which HHI was calculated. Localities were determined through K-means clustering from point address locations of completed heat pump installations across Massachusetts.

All submarkets sit below the 2,500 mark, informing us of a reasonable enough level of competition in each locality for consumers to have a reasonable choice across competitors. The location with the highest amount of market concentration is locality 13, with an HHI of 2,411, which is in the far west of Massachusetts and has the lowest density of installations. The lowest amount of concentration is seen in locality 2, in the Boston area, with an HHI 306. Low market concentration is not exclusive to the Boston area. However, areas such as localities 15 and 4 are in less dense areas of Massachusetts and still have high levels of market competition.

4 Empirical Strategy

The estimation of the impact of contractor installation experience on the total cost of the project and installed heating capacity is undertaken through a fixed effects modeling approach. The models control for relevant characteristics of installations, such as the hedonic attributes of the home, as well as the fixed effect estimates of individual contractors, to estimate the effect of added experience (learning) on installation outcomes.

4.1 Linear Model for Learning

We use a fixed-effects model to estimate the relationship between added experience in heat pump installations and installation outcomes, conditional on the relevant characteristics of the home and heating system:

$$ln(Y_{i,c,t}) = \lambda_c \cdot + \tau_t + \gamma \cdot ln(Experience_{c,t}) + \beta X_i + \epsilon_{i,c,t} \quad (1)$$

$Y_{i,c,t}$ describes the two main outcomes describing the performance and costs of each installation in our sample: $ln(Capacity/Sqft_{i,c,t})$, and $ln(TotalCost/Sqft_{i,c,t})$ represent the log of installed capacity (per square footage) of the heating system and the total cost (per square footage) of the project for home i completed by contracting company c in time t. λ_c describes the fixed effects associated with contracting company c, and τ_t describes the time fixed effects. Our coefficient of interest is γ, which describes the changes in costs and installed capacity for each additional heat pump installation by contractor c at year t.

Lastly, β includes a vector of coefficients associated with the list of controls for characteristics of the heating system and home that are controlled for. In particular, the vector includes the size, year built, value of the property, type of house, number of bedrooms and bathrooms, and the baseline heating type in the property at the time of the heat pump installation. The standard errors are clustered at the contracting company level (c), for the purpose of capturing unobservables correlated at the contractor level.

5 Results

5.1 Impact of Learning on Contractors

Table 2 panel A presents the association between contractor experience and the heat-pump capacity installed in the house. The table presents our main estimates in several model specifications varying the controls and fixed effects described in Eq. 1. Once controlling for time effects, we estimate that added experience leads to a decrease in installed capacity. As controls for the characteristics of the home and location enter the model, this effect becomes increasingly significant and negative. For our final model (6), the interpretation for the estimates is that for each additional percentage increase in installations completed, the next heat pump the contractor installs is downsized by -0.033%.

Table 2: Relationship between installer experience, installed capacity, and installation costs.

	(1)	*(2)*	*(3)*	*(4)*	*(5)*
	(A) Installed Capacity Models				
Ln(Number Past	-0.026**	-0.029**	-0.031**	-0.033***	-0.033***
Installations)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
DF Residuals	13,503	13,017	12,992	12,990	12,990
R-Squared	0.27	0.33	0.36	0.36	0.36
	(B) Total Cost of Installation Models				
Ln(Number Past	-0.011	-0.018	-0.020	-0.021	0.011
Installations)	(0.013)	(0.014)	(0.014)	(0.014)	(0.0072)
Ln(Installed Capacity)	-	-	-	-	0.94*** (0.0093)
DF Residuals	13,503	13,017	12,992	12,990	12,989
R-Squared	0.34	0.38	0.4	0.4	0.84
	Controls				
Installer FEs	Yes	Yes	Yes	Yes	Yes
Year Installed FEs	Yes	Yes	Yes	Yes	Yes
Month Installed FEs	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes
Town FEs	No	Yes	Yes	Yes	Yes
Previous Heating Controls	No	No	Yes	Yes	Yes
Heat Pump Brand	No	No	No	Yes	Yes

Table 2 panel B presents the models that examine the relationship between the total cost of installation as the dependent variable and increased installer experience. Once controlling for time effects, the estimate for the effect of additional installation is negative, but consistently remains non-significant, with relatively large standard errors throughout various sets of controls introduced.

5.2 Magnitude of Learning Effect at Different Levels

The ”previous installations” variable in the standard regression is counted on the contracting company level. The possibility of learning occurring on a more granular level (by individual

Table 3: Relationship between experience and installed capacity at different aggregation levels.

	(1)†	*(2)†*	*(3)†*
	Company Level	*Individual Level*	*Town Level*
Ln(Number Past	-0.040***	-0.026***	-0.035***
Installations)	(0.0095)	(0.0090)	(0.013)
Observations	12,824	12,286	13,456
R-Squared	0.56	0.58	0.52

†Controlling for installer, time, town, previous heating source, heat pump type, and hedonics.

$***p < 0.01, **p < 0.05, *p < 0.1$

employee) and on a more macro level (by town) is shown in Table 3. In all three cases, the effect of additional installations on installed heat pump capacity is negative and significant. The magnitude of this effect remains the greatest at the company level, as used in the rest of the study. This provides an important implication that the potential heat pump installation training programs that target the company level may be the most effective approach.

5.3 Heterogeneity Analysis

Separate regressions are made for heat pump types (single/multi-head) in table 4 and previous fuel sources (natural gas, oil, other) in table 5. For installed capacity, the relationship with additional experience is most pronounced for single-head heat pumps, and also the retrofits that replace heating systems that were primarily fueled by natural gas. In the case of installation costs, the relationship with experience when including all categories is insignificant, but becomes significant for the subset of single-head heat pumps, and also the subset of homes previously with gas-based heating systems.

To understand the heterogeneity patterns we observe above, we interviewed major heat pump installers in Massachusetts. From such interviews, we have learned that the complexity related to installing single-head or multi-head heat pumps, and natural gas or oil heater replacements may give some insights into such variations.

The available range of a heat pump is measured by the ”turn-down ratio”, defined as the ratio of the maximum to the minimum capacity for a targeted indoor temperature. For a comparison between installing either one or multiple single-head heat pumps and one multi-head heat pump for a home, the consideration of turn-down ratios for a multi-head heat pump is a more complex task. Table 4 shows that the learning effect is more salient for single-head heat pump installations as a relatively less complex task.

The different levels of task complexity related to replacing the previously gas-based or oil-based heating systems with heat pumps can also help explain the learning effect heterogeneity

Table 4: Relationship between installer experience, installed capacity and installation costs, by heat pump type.

	(1)†	*(2)†*	*(3)†*
	All Types	*Single-Head*	*Multi-Head*
(A) Outcome: Installed Capacity			
Ln(Number Past Installations)	-0.031***	-0.055***	-0.023
	(0.012)	(0.013)	(0.017)
Observations	12825	5847	6179
R-Squared	0.36	0.77	0.83
(B) Outcome: Installation Cost			
Ln(Number Past Installations)	-0.021	-0.0544***	-0.0089
	(0.014)	(0.015)	(0.018)
Observations	12825	5847	6179
R-Squared	0.41	0.82	0.86

†Controlling for installer, time, town, previous heating source, heat pump type, and hedonics.

$***p < 0.01, **p < 0.05, *p < 0.1$

we see in Table 5. Gas and oil furnaces, by nature, create heat through combustion and heat air at a much higher temperature than a heat pump does (with all three being equally capable of heating a home). Oil furnaces, however, typically burn hottest and thus use smaller, high-velocity air ducts for properly distributing air at a safe temperature. When installing heat pumps, these small ductwork would be grossly undersized for the new heat pump put in place, leading to air duct leakage and poor performance. Therefore, a more complex task of replacing ductwork is always required for the homes previously heated by oil, which may lead to a smaller learning effect we observe here, compared to those homes previously heated by natural gas.

6 Conclusion and Next Steps

In this paper, we examine the presence of learning effects in the installations of heat pumps. By creating a unique data set that combines the details of heat pump installation and contractor companies with a rich set of home characteristics, we estimate the relationship between installation experience, measured by the number of heat pump installations a contractor has completed, and the sizing of the heat pump along with the total installation costs.

With these data, we first identify large amounts of heterogeneity in heat pump sizes and installation costs across contractor companies after controlling for the characteristics of heat pumps and homes. This high variance in cost and heating capacity not only leads to a greater

Table 5: Relationship between installer experience, installed capacity and installation costs, by previous heating fuel.

	(1)†	*(2)†*	*(3)†*	*(4)†*
	All Types	*Natural Gas*	*Oil*	*Other*
(A) Outcome: Installed Capacity				
Ln(Number Past Installations)	-0.031***	-0.045**	-0.026	-0.014
	(0.012)	(0.021)	(0.016)	(0.031)
Observations	12825	4409	5001	2025
R-Squared	0.36	0.79	0.76	0.90
(B) Outcome: Installation Cost				
Ln(Number Past Installations)	-0.021	-0.046**	-0.0076	-0.0081
	(0.014)	(0.021)	(0.020)	(0.033)
Observations	12825	4409	5001	2025
R-Squared	0.41	0.80	0.78	0.91

†Controlling for installer, time, town, previous heating source, heat pump type, and hedonics.

$***p < 0.01, **p < 0.05, *p < 0.1$

uncertainty and a big barrier for households to understand the upfront costs of an installation and its benefits but also indicates the existence of sub-optimal choices of heat pumps, potentially undermining the goal of the heat pump rebate program.

We find a consistently significant and negative association between increased past installations and the installed capacity of heating systems. We also find that the sizing of the heat pump explains much of the total costs of heat pump installation, indicating that an effective learning process can help correct the over-sizing problem, thus reducing installation costs and improving energy efficiency. Our results also indicate that this learning effect is more salient for less complex installation tasks such as single-head heat pumps and in those homes previously using natural gas heating systems.

Our next step is to explore how such a learning effect of contractors affects the final electricity consumption of households. In addition, we are currently implementing energy models in our sample of houses to examine what the optimal installed capacity should be, how the selected capacity by the installer deviates from this optimal capacity, and how the learning experience can help them converge to it.

References

Allcott, H. and M. Greenstone (2012): "Is There an Energy Efficiency Gap?" Journal of Economic Perspectives, 26, 3–28.

Bollinger, B. and K. Gillingham (2019): "Learning-by-Doing in Solar Photovoltaic Installations," .

Christensen, P., P. Francisco, E. Myers, and M. Souza (2021): "Decomposing the Wedge between Projected and Realized Returns in Energy Efficiency Programs," The Review of Economics and Statistics, 1–46.

Domanski, P. A., H. I. Henderson, and W. V. Payne (2014): "Sensitivity Analysis of Installation Faults on Heat Pump Performance," Tech. Rep. NIST TN 1848, National Institute of Standards and Technology.

Dun & Bradstreet (2023): "Zillow Group, Inc. Company Profile: Competitors, Financials & Contacts," .

European Comission (2020): "Stepping up Europe's 2030 climate ambition, Investing in a climate-neutral future for the benefit of our people," Tech. Rep. Brussels, 17.9.2020, European Comission.

Gindelsky, M., J. G. Moulton, and S. A. Wentland (2020): "Valuing Housing Services in the Era of Big Data: A User Cost Approach Leveraging Zillow Microdata," in Big Data for Twenty-First-Century Economic Statistics, University of Chicago Press, 339–370.

Gleeson, C. P. (2016): "Residential heat pump installations: the role of vocational education and training," Building Research & Information, 44, 394–406.

Heidari, M., S. Zad, and S. Rafatirad (2021): "Ensemble of Supervised and Unsupervised Learning Models to Predict a Profitable Business Decision," in 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 1–6.

Holt, J. R. and M. E. Borsuk (2020): "Using Zillow data to value green space amenities at the neighborhood scale," Urban Forestry & Urban Greening, 56, 126794.

International Energy Agency (2022a): The Future of Heat Pumps, OECD.

——— (2022b): "Heating - Subsector Report," .

——— (2022c): "Technology Report: Installation of about 600 million heat pumps covering 20% of buildings heating needs required by 2030," .

Jovanovic, B., Y. Nyarko, and Griliches (1995): "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves," Brookings Papers on Economic Activity. Microeconomics, 1995, 247–305, publisher: Brookings Institution Press.

Kerry, J. (2021): "The Long-Term Strategy of the United States, Pathways to Net-Zero Greenhouse Gas Emissions by 2050," Tech. rep.

Mass.gov (2023a): "Apply for an Oil Burner Technician or Apprentice Certificate," .

——— (2023b): "Board of State Examiners of Plumbers and Gas Fitters," .

Thompson, P. (2010): "Chapter 10 - Learning by Doing," in Handbook of the Economics of Innovation, ed. by B. H. Hall and N. Rosenberg, North-Holland, vol. 1 of Handbook of The Economics of Innovation, Vol. 1, 429–476.

US Department of Justice (2015): "Herfindahl-Hirschman Index," .

US EPA, O. (2015): "Section 608 Technician Certification," .

7 Appendix

7.1 Data in Detail

7.1.1 Air-Source Heat Pump Residential Projects Database

These data represent individual installations of heat pumps in homes. When applying for the rebate, information about the home and installation was recorded for determining eligibility, and the recorded values were compiled into this data set summarized in *table 6*. For each installation, the total cost of the installation was recorded along with the identity of the contractor, allowing us to observe the set of homes each contractor worked on. The address data of the home was additionally used to match additional data of homes such as with the Zillow data.

Date of Installation	
Total Cost	The total recorded cost of installation, including equipment and labor.
Square Footage	The square footage of living area in the home.
County	The Massachusetts county in which the home is located.
Town	The town in which the home is located.
Past Fuel	The heating fuel source used by the home (i.e. gas or oil).
Num. Units	The number of heat pump units installed.
Installed Capacity	The total installed heating capacity of the installed heating system (BTU/Hr).
Is Retrofit	Whether installation was in existing or home completing construction.
Home Address	The address of the home.
Contractor	The name and address of the contractor who worked on the project.

Table 6: Variables Available in ASHP Rebate Data

7.1.2 Zillow Data

The postal address of the homes in the ASHP dataset were used to search and scrape information off of the Zillow website. The variables collected, when available, are listed in *table 7*. The importance of collecting data on the characteristics of the home and type of installation is to use the variables as additional controls for the fixed effects analysis. Not including a variable for the year of construction, for example, could lead to the observed variation between contractor effects be due to the unobserved variability in a home's age. Data made publicly available by Zillow has been similarly used in other research for gaining parcel-level information on properties and the neighborhoods they reside in Gindelsky et al. (2020); Holt and Borsuk (2020); Heidari et al. (2021).

Parcel Number	Uniquely identifying number of home.
Home Type	The type of home (i.e. Single Family or Multi Family).
Roof Type	The type of roofing installed (i.e. Asphalt).
Exterior Features	Primarily the exterior finish of the home (i.e. Shingle or Brick).
Construction Materials	Material used in construction (i.e. Frame or Concrete).
Living Area	Square footage of available living area.
Bedrooms	Number of bedrooms.
Bathrooms	Number of bathrooms.
Year Built	The recorded year of completion of the home.
Heating	Most recent recorded heating method of home.
Cooling	Most recent recorded cooling method of home.
Price History	Recorded prices of home when listed on the market as available to Zillow.
Tax History	Recorded tax history of home as available to Zillow.
Zestimate	Zillow's current estimate of the home's market value.

Table 7: Variables Scraped from Zillow

Data Cleaning and Transformation

Heat Pump Related Variables

The primary characteristic of the heat pump system that is available is its installed capacity. For the regressions the natural log of the installed capacity is taken.

Home Characteristics Variables

- Footage and living area: there are two sources of information for the size of the home, the square footage of the living area provided by the survey data, and the footage reported by Zillow. The footage reported in the survey data is used unless in the case where it is too small (below 1000), and the other value is a more plausible alternative.

- Value of the home: The value of the home is produced from the reported valuation generated by Zillow, called the "Zestimate". For the regressions the natural log of the Zestimate is taken to satisfy assumptions of normality.

- Age of house: The year in which the home was built is sourced from the home's Zillow page. The structure's age is calculated as the present year (2023) minus the year of completion.

- Bedrooms and Bathrooms: Both variables are sourced from Zillow. Homes with less than one bedroom and bathroom are excluded from the data, and values above 5 are truncated at 5. The bedroom and bathroom variables are ultimately interpreted as categorical variables.

Contractor Variables

- Contracting Company Identity: For each heat pump installation recorded the name of the contracting company that completed it, along with a phone number, and address of the business is recorded. Contractors for the purposes of the model are identified through name after being de-duplicated by removing casing and symbols.

- Previous Installations: The heat pump installation data is longitudinal, and at each new installation, the number of previous installations completed by a contractor within the data set are summed. For the regressions the natural log of this variable is taken.