HETEROGENEOUS EFFECTS OF COOLING DEMAND ON HOUSEHOLD ELECTRICITY CONSUMPTION USING CAUSAL ML MODELS

Tosin Kolajo Gbadegesin

Ph.D. Economics Candidate

Primary Fields: Energy and Environmental Economics

Secondary Fields: Urban Economics

Department of Economics,

Howard University, Washington DC

tosin.gbadegesin@bison.howard.edu | tosingbade05@gmail.com

ABSTRACT

This study examines the heterogeneous causal effects of climate-driven cooling demand on U.S. household electricity consumption using 2020 Residential Energy Consumption Survey data and a causal machine learning framework. Estimates of conditional average treatment effect (CATE), conditional quantile treatment effect (CQTE), and conditional super quantile treatment effect (CSQTE) show substantial increases in electricity consumption under elevated cooling degree days (CDD), with average effects of 32–34% and disproportionate burdens concentrated among lower-consuming households. OLS projections of CATE, CQTE, and CSQTE estimates highlight significant subgroup variation: households with central AC and evaporative coolers, middle-income earners ($60k–$79k), and minority racial groups face the highest burdens, while single-family homes exhibit consistently stronger responses than apartments or mobile homes. Robustness checks using alternative treatment definitions (CDD above the 75th percentile and 30-year average CDD) confirm the stability of results. These findings underscore that climate-driven cooling demand exacerbates distributional inequalities in energy use, reinforcing the need for targeted efficiency and adaptation policies that account for subgroup vulnerabilities.

Keywords: Climate change; Cooling demand; Household electricity consumption; Causal machine learning; Treatment effect heterogeneity; Distributional effects

1.0 Introduction

Rising heat exposure under climate change is reshaping residential electricity use in uneven ways across households and regions, with important implications for affordability, peak-load management, and equity (Hernández 2013; Carley and Konisky 2020; Bednar and Reames 2020). In the United States, cooling needs have become an increasingly salient driver of household electricity demand as heatwaves intensify and lengthen, altering seasonal load profiles and stressing distribution networks (Steinberg et al. 2020; Bawaneh et al. 2024). Cooling degree days (CDD), widely used in studies such as Sailor (2001) and Isaac and van Vuuren (2009) to quantify temperature-driven cooling demand provide a relevant metric for modeling temperature–electricity relationships.

A large literature documents strong links between temperature and residential electricity consumption, including sharp responses during extreme heat events. For example, Deschênes and Greenstone (2011) show that temperature fluctuations significantly affect residential energy demand in the United States, while Auffhammer and Aroonruengsawat (2011) demonstrate pronounced nonlinear responses of electricity consumption to high temperatures, particularly during heat extremes. However, much of this evidence relies on average effects, which can mask substantial heterogeneity in how households respond to rising cooling demand. Evidence from pricing experiments and demand-response programs further shows wide dispersion in behavioral responses across households with different technologies, housing characteristics, and resources (Reiss and White 2005; Ito 2014; Knittel and Stolper 2022). These findings highlight the need for empirical approaches that can identify not only average impacts, but also how climate-driven cooling demand affects households differently across the consumption distribution.

This study addresses three related questions. First, what is the causal effect of increased cooling demand on household electricity consumption? Second, how do these causal effects vary across households with different socioeconomic and housing characteristics? Third, how are these effects distributed across the electricity consumption distribution, particularly among households with high baseline electricity use? Together, these questions aim to identify which households bear the greatest marginal electricity demand response to rising cooling needs.

To answer these questions, the chapter employs causal machine learning methods. Unlike traditional regression approaches that impose linearity and low-dimensional interaction structures, causal machine learning allows treatment effects to vary flexibly with high-dimensional characteristics (Athey et al. 2019; Chernozhukov et al. 2018; Wager and Athey 2018). By capturing nonlinear interactions while employing orthogonalization and cross-fitting to reduce bias and overfitting (Chernozhukov et al. 2018; Nie and Wager 2021; Künzel et al. 2019), these methods support valid inference under standard identification assumptions. The analysis estimates conditional average treatment effects (CATE) as well as conditional quantile and super-quantile treatment effects (CQTE and CSQTE), enabling a joint examination of heterogeneity across households and across the consumption distribution. Estimated treatment effects are subsequently projected onto observable household and housing characteristics to enhance interpretability and clarify the sources of differential vulnerability to cooling demand.

The analysis uses 2020 RECS microdata and exploits within–climate-zone variation in cooling exposure combined with rich household covariates to support a causal interpretation of heterogeneous treatment effects. The resulting estimates provide new causal evidence on how climate-driven cooling demand differentially shapes household electricity consumption across the distribution, offering policy-relevant insights for the design of targeted, equitable, and adaptive energy and climate policies.

The remainder of this study is organized as follows. Section 2 presents the literature review, summarizing prior evidence on causal ML approaches and highlighting the contribution of this study to the growing causal and distributional literature on climate–energy interactions. Section 3 outlines the methodological framework, including data description, treatment construction, and the empirical model. Section 4 discusses the empirical results, emphasizing heterogeneity across household groups and consumption levels. Finally, Section 5 concludes with key policy implications for energy equity and climate adaptation.

2.0 Literature Review

2.1 Conditional Average Treatment Effects Estimation

The estimation of Conditional Average Treatment Effects (CATE) is central to understanding treatment effect heterogeneity, particularly in observational and experimental studies. Over the

years, researchers have advanced various methodologies employing statistical and machine learning techniques to address challenges such as high-dimensional data, confounding variables, and treatment effect heterogeneity. Hill (2011) introduced Bayesian Additive Regression Trees (BART), a nonparametric Bayesian method for modeling response surfaces. BART flexibly estimates heterogeneous treatment effects by focusing on the response surface, outperforming traditional methods like propensity score matching in nonlinear settings. Its ability to handle high-dimensional covariates makes it particularly suitable for complex datasets. Similarly, Souto and Neto (2024) expanded on BART with their K-Fold Causal BART model, which combines Bayesian regression trees with cross-validation to improve the estimation of both ATE and CATE, though it exhibited limitations in certain real-world datasets.

Strittmatter (2023) demonstrated the practical utility of causal machine learning (CML) in policy evaluations, where CML offered nuanced insights into treatment heterogeneity beyond traditional estimators. Kim et al. (2023) further emphasized the importance of CATE estimation in clustered data, applying methods such as causal forests and BART to address cross-level interactions. These methods provided valuable insights into educational policy decisions by accounting for the interactions between individual and community-level covariates. Kato and Imaizumi (2023) introduced CATE Lasso, which leverages implicit sparsity in high-dimensional linear models to achieve robust and consistent estimates, validating its effectiveness in distinguishing treatment-specific effects through simulation studies.

Hahn et al. (2020) proposed Bayesian causal forests as an advanced method for estimating heterogeneous treatment effects in the presence of strong confounding. By incorporating propensity score estimation and separately regularizing treatment heterogeneity, their approach addressed biases in nonlinear models. Similarly, Gbadegesin and Yameogo (2024) demonstrated the potential of hybrid methodologies by combining machine learning models like Gradient Boosting and XGBoost with targeted maximum likelihood estimation. Their work highlights how updating initial predictions with clever covariates improves treatment effect estimation efficiency.

Wager and Athey (2018) introduced the causal forest algorithm, which extends Breiman's random forests to estimate heterogeneous treatment effects. Their method is notable for its ability to construct asymptotically valid confidence intervals, making it particularly effective in high-

dimensional settings. Expanding on causal forests, Oprescu et al. (2019) proposed orthogonal random forests, which integrate Neyman-orthogonality with generalized random forests. This approach reduces sensitivity to nuisance parameter estimation errors and performs robustly in both discrete and continuous treatment scenarios.

Künzel et al. (2019) developed meta-algorithms, including the X-learner, which enhances the efficiency of CATE estimation, especially when treatment group sizes are imbalanced. The X-learner integrates base learners like random forests and BART, leveraging structural properties of the CATE function for flexible and efficient estimation. Extending this line of research, Machluf et al. (2024) introduced ensemble methods such as the Stacked X-Learner and Consensus-Based Averaging (CBA). These methods aggregate multiple estimators to improve robustness and stability in CATE estimation, particularly in scenarios with high uncertainty or varying data-generating processes, demonstrating superior performance across diverse datasets.

2.2 Double Machine Learning and Orthogonal Statistical Learning

Double machine learning (DML) and orthogonal statistical learning have emerged as robust methodologies for addressing causal inference challenges, particularly in high-dimensional settings with complex nuisance parameters. These methods leverage advancements in machine learning to mitigate issues like bias from regularization and overfitting while enabling efficient estimation of treatment effects, including heterogeneous effects. Chernozhukov et al. (2018) introduced the DML framework, which uses Neyman-orthogonal moments and cross-fitting to achieve consistent and asymptotically normal estimates of CATE. This framework is highly flexible, accommodating machine learning algorithms like random forests, lasso, ridge regression, and neural networks for nuisance parameter estimation. The versatility of DML has been demonstrated in models ranging from partially linear regression to instrumental variables and treatment effect estimation, underscoring its broad applicability.

Building on Chernozhukov et al. (2018), Bach et al. (2022) developed DoubleML, an open-source Python library that integrates DML with machine learning models, facilitating its adoption in empirical research. This tool provides functionalities for valid statistical inference across various causal models, solidifying DML's position as a cornerstone in modern causal inference. Knaus (2021) showcased DML's adaptability by applying it to assess the effects of musical practice on

youth development. This study highlighted practical considerations in implementing DML, such as parameter tuning and covariate balancing, demonstrating its ability to address real-world complexities while maintaining robust inference.

Orthogonal learning has also made significant contributions to causal inference, particularly through its capacity to manage nuisance parameter estimation errors. Foster and Syrgkanis (2019) provided theoretical guarantees for excess risk in models involving nuisance parameters, demonstrating the robustness of orthogonal learning in high-dimensional and nonparametric settings. Mackey et al. (2018) and Zadik et al. (2018) extended this framework to orthogonal machine learning, introducing higher-order orthogonality to enhance robustness in non-linear and complex causal inference tasks. Athey and Wager (2021) further advanced the field by applying orthogonal statistical learning to policy optimization. Their methodology integrates doubly robust estimators with optimization techniques, allowing for the design of treatment policies that balance fairness, budget constraints, and other considerations.

Extensions of the DML framework to heterogeneous treatment effect estimation have provided critical insights into optimality in causal inference. Kennedy (2020) proposed a doubly robust CATE estimator that achieves faster error rates in smooth nonparametric models, advancing the understanding of statistical limits in heterogeneous treatment effects. Ichimura and Newey (2022) emphasized the utility of influence functions in semiparametric models for debiasing estimators, providing robust tools for policy evaluation in the presence of complex nuisance structures. These advancements have proven transformative for empirical research, as demonstrated by applications in dose-response relationships (Knaus, 2021) and policy learning (Athey & Wager, 2021).

Despite their strengths, DML and orthogonal learning methods face practical challenges, including computational complexity and sensitivity to tuning parameters. Kennedy (2020) and Knaus (2021) highlighted the importance of careful parameter tuning and covariate balancing in high-dimensional applications. Additionally, Mackey et al. (2018) and Zadik et al. (2018) noted the scalability limitations of higher-order orthogonal moments in large datasets. These challenges underscore the need for further methodological refinements to enhance the scalability and efficiency of these approaches in applied settings.

2.3 Distributional Treatment Effects Estimation

The estimation of distributional treatment effects (DTE) provides a comprehensive framework for understanding how treatments impact the entire distribution of an outcome, rather than limiting the analysis to average effects. Chernozhukov et al. (2013) laid the foundation for modeling counterfactual distributions using regression-based methods, enabling the analysis of full outcome distributions. Their work introduced techniques to construct confidence sets for functional effects like quantile functions, facilitating applications in policy analysis such as wage decompositions. This framework offered distribution regression as a robust alternative to quantile regression, emphasizing its flexibility. Later, Chernozhukov et al. (2020) extended these methodologies to nonlinear network and panel models, addressing challenges like unobserved two-way effects and incidental parameter problems. Their approach, which constructs confidence bands for quantile functions, has proven effective in analyzing complex datasets, including trade networks.

The integration of machine learning into DTE estimation has significantly advanced the field, offering new tools for addressing treatment heterogeneity. Zhou and Carlson (2021) proposed Collaborating Causal Networks (CCN), a flexible framework for estimating full potential outcome distributions without restrictive assumptions about the data-generating process. CCN excels in capturing distributional variations, particularly in tail behavior and risk profiles, making it a valuable tool for nuanced decision-making. Park et al. (2021) introduced the Conditional Distributional Treatment Effect (CoDiTE), which generalizes the CATE framework to higher-order moments of treatment effects. Their methodology, based on kernel conditional mean embeddings and U-statistic regression, has been successful in analyzing real-world and synthetic datasets, highlighting its utility in understanding tail events and distributional shifts. Kallus and Oprescu (2023) further contributed by developing a model-agnostic approach for estimating Conditional DTEs (CDTEs). Their pseudo-outcome-based framework provides robust estimates of quantile and super-quantile treatment effects, particularly under model misspecification, making it valuable for fields like financial risk assessment.

Machine learning has also enhanced the precision of DTE estimation in experimental settings. Byambadalai et al. (2024) developed a regression adjustment method that integrates pre-treatment covariates into distributional regression frameworks, reducing variance and improving estimator precision in randomized experiments. Linden and Yarnold (2016) tackled the challenges of

multivalued treatments by employing optimal discriminant analysis, which they argued outperforms regression-based estimators in capturing heterogeneity across treatment groups. Similarly, localized debiased machine learning, introduced by Kallus et al. (2019), addresses complexities in quantile treatment effect estimation by focusing on parameter-dependent nuisances, achieving robust performance in high-dimensional settings. Belloni et al. (2017) further emphasized the importance of orthogonal moment conditions for ensuring valid inference in high-dimensional environments, demonstrating their utility in estimating both local and global QTEs.

Recent innovations have also highlighted the scalability of DTE estimation in applied contexts. Wu et al. (2023) introduced DNet, a novel architecture capable of estimating distributional individualized treatment effects. DNet's strength lies in its ability to model entire outcome distributions, particularly in heavy-tailed settings, and its successful deployment in applications like mobile app optimization underscores its robustness and scalability. Finally, Curth et al. (2024) explored the application of DTEs in clinical settings, focusing on tailoring treatments to individual patient characteristics. Their review emphasized the challenges of covariate shift and identification assumptions, advocating for methodological advancements to enhance the applicability of DTE estimation in diverse contexts.

Hsu et al. (2023) introduced a double/debiased machine learning framework to estimate direct and indirect quantile treatment effects, achieving robust results even under model misspecifications. Their method was validated through simulations and applied to assess earnings effects in the national job corps study. Kallus et al. (2024) proposed localized debiased machine learning to efficiently estimate local quantile treatment effects while reducing computational complexity, demonstrating robust performance in high-dimensional settings. Similarly, Chen, Huang, and Tien (2021) extended debiased machine learning to instrumental variable quantile regression, effectively handling high-dimensional controls and providing insights into the quantile treatment effects of 401(k) participation on wealth. These studies underscore the growing precision and applicability of DTE estimation techniques.

2.4 Application of CATE and DTE in Energy and Climate Research

While the application of ML to CATE and DTE is well-established in fields such as healthcare and economics, its integration into energy and climate research is still emerging. Currently,

relatively few studies explicitly apply ML methods to estimate CATE and DTE in these fields. For example, Knittel and Stolper (2021) applied causal forests to assess the heterogeneous effects of energy conservation nudges, uncovering substantial variations in household electricity reductions and highlighting the potential for ML to personalize energy interventions. Klosin and Vilgalys (2022) employed DML to estimate the effects of extreme heat on U.S. corn production, showcasing the effectiveness of ML methods in capturing complex, non-linear relationships and quantifying the impacts of extreme weather. In climate-focused applications, Giannarakis et al. (2022) employed CATE estimation to assess the effects of sustainable agricultural practices on soil organic carbon, emphasizing the need for localized interventions to enhance carbon sinks. Gadea and Gonzalo (2023) advanced the analysis of climate heterogeneity by examining the full temperature distribution, and regional variations in warming patterns that demand tailored mitigation strategies.

This study applies the robust, model-agnostic framework for CDTE estimation developed by Kallus and Oprescu (2023), alongside CATE estimation, to address key challenges in energy and climate research, including model misspecification, high-dimensional covariates, and the need to capture heterogeneity across the outcome distribution. While CATE estimation identifies average treatment effect variation across subgroups, CDTE extends this by quantifying how treatment effects differ at different points of the conditional outcome distribution, providing a richer picture of heterogeneity. Both approaches leverage recent advances in causal machine learning, including pseudo-outcome regressions, cross-fitting, orthogonalization, and flexible learners such as causal forests and double/debiased machine learning, ensuring robustness without relying on restrictive functional form assumptions.

3.0 Methodology

3.1 Data Sources and Description

This study uses data from the 2020 Residential Energy Consumption Survey (RECS), a nationally representative dataset collected by the U.S. Energy Information Administration (EIA) that provides detailed information on household energy use, housing characteristics, and demographics across the United States.

The dataset undergoes systematic cleaning prior to analysis. Observations with imputed values for key variables like electricity use and thermostat settings are excluded using EIA-provided imputation flags. Continuous variables are trimmed at the 1st and 99th percentiles to reduce outlier influence and log-transformed to address right-skewed distributions, enabling elasticity-based interpretations. Categorical variables, including appliance ownership, household type, and demographic characteristics, are retained in their original form unless invalid or missing. After cleaning, the final analytical sample includes 13,198 households from 18,496 households. Descriptive statistics are weighted to ensure national representativeness, with means, proportions, and percentiles reflecting population-level distributions.

3.2 Descriptive Statistics

Table 1 summarizes the composition of U.S. households in the 2020 RECS sample based on key categorical characteristics. Geographically, households are concentrated in the Cold & Very Cold (35%) and Mixed Humid (34%) climate zones, with smaller shares located in the Hot Humid (18%), Hot/Mixed Dry (10%), and Marine (2%) regions. In terms of cooling systems, the majority of households (78%) rely on central air conditioning, while 20% use unitary AC systems, and only 2% utilize evaporative coolers. The income composition shows that over a quarter (28%) of households earn more than $100,000 annually, while 13% fall below the $20,000 threshold. The sample is predominantly composed of White households (83%), followed by Black (10%), Asian (4%), and smaller shares of multiracial, Native American, and Pacific Islander households. In terms of housing structure, 71% of households live in single-family homes, with apartments (24%) and mobile homes (6%) comprising the rest. Finally, 82% of households report the presence of insulation, which plays a significant role in moderating cooling energy demand.

Table 1: Weighted proportions of categorical variables

	Unweighted Count	Weighted Population	Weighted Proportion
Marine Climate Zone	306	2,076,919	0.02
Cold & Very Cold Climate Zone	5407	31,731,365	0.35
Hot Humid Climate Zone	1780	16,583,898	0.18
Hot and Mixed Dry Climate Zone	985	8,655,676	0.10
Mixed Humid Climate Zone	4720	30,732,803	0.34
Central AC Type	10329	69,848,961	0.78
Unitary AC Type	2519	17,983,154	0.20
Evaporative Coolers	350	1,948,547	0.02
Below 20k	1492	11,879,551	0.13
20k-39k income group	2333	17,340,342	0.19
40k-59k income group	2105	14,757,166	0.16
60k-79k income group	1458	9,611,087	0.11
80k-99k income group	1758	11,239,435	0.13
100k and above income group	4052	24,953,081	0.28
Asian Households	452	3,584,970	0.04
Black Households	1083	8,930,101	0.10
Multiracial Households	276	2,107,728	0.02
Native American Households	92	629,053	0.01
Pacific Islander Households	21	176,589	0.00
White Households	11274	74,352,220	0.83
Apartment Homes	2235	21,158,557	0.24
Mobile Homes	699	5,020,957	0.06
Single Family Homes	10264	63,601,148	0.71
Presence of Insulation	11112	73,646,759	0.82
Not Insulated	2086	16,133,903	0.18

Table 2 presents weighted descriptive statistics for key continuous variables used in the analysis. On average, U.S. households consumed approximately 2,153 kWh of electricity annually for cooling, with substantial variation across the sample (SD: 1,702 kWh), and values ranging from 76 kWh to over 10,000 kWh. Climate exposure measures show an average of 1,674 CDD and 3,787 HDD. The typical household cooled about 1,484 square feet, lived in a dwelling with six rooms, and had an average household size of two people. In terms of long-term climate norms, the 30-year CDD average (CDD30YR) was 1,449, with a right-skewed distribution extending to nearly 4,000. Lighting behavior, measured by the number of bulbs on for 1–4 hours daily, had a mean of 5 bulbs, with notable variation across households (ranging from 0 to 30).

Table 2: Weighted statistics for continuous variables

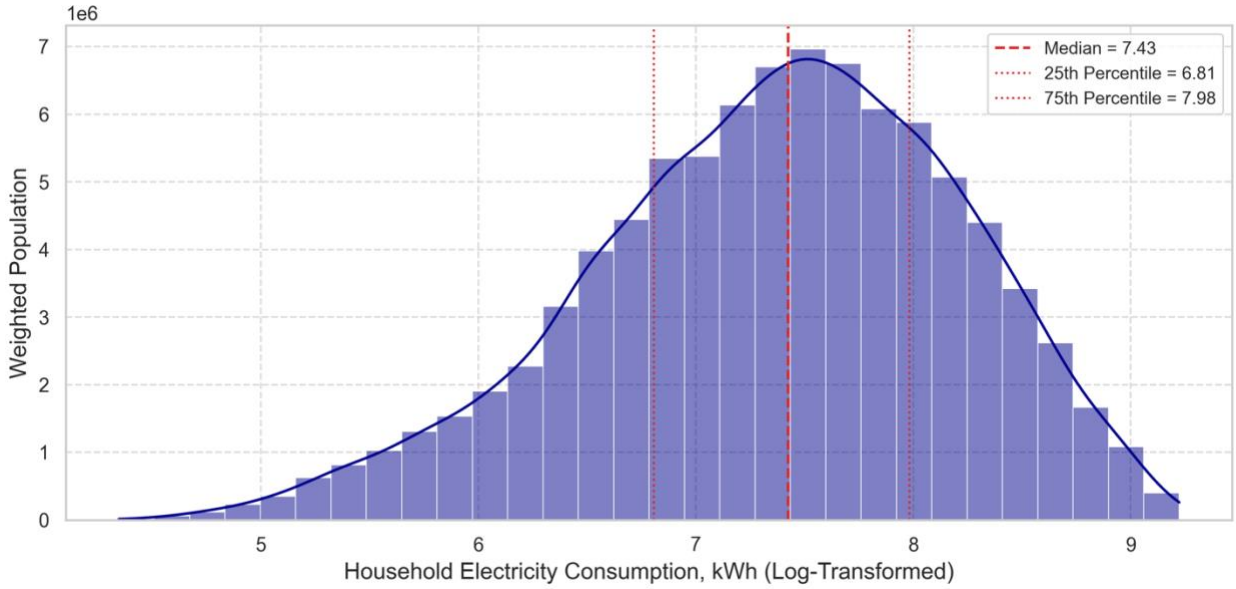
	Count	Mean	Std Dev	Min	25th Percentile	Median	75th Percentile	Max
Household Electricity Consumption (kWh)	13198	2152.83	1701.92	76.32	903.22	1677.49	2923.20	10124.60
Cooling degree days	13198	1674.06	1008.05	221.00	954.00	1335.00	2067.00	5082.00
Heating degree days	13198	3786.52	1981.50	99.00	2083.80	4172.74	5356.00	8681.00
Cooled SqFt	13198	1483.98	939.43	140.00	780.00	1300.00	2000.00	4800.00
Total Number of Rooms	13198	6.33	2.25	1.00	5.00	6.00	8.00	15.00
Household Size	13198	2.43	1.35	1.00	1.00	2.00	3.00	7.00
CDD30YR	13198	1449.35	872.17	209.00	781.00	1186.00	1898.00	3965.00
Bulbs on for 1-4hrs	13198	5.08	4.90	0.00	2.00	4.00	6.00	30.00

3.3 Outcome Variable of Interest

The primary outcome variable is annual household electricity consumption for space cooling, measured in kilowatt-hours (kWh). This measure is derived from calibrated RECS estimates, reflects household-level cooling loads across different system types and climate conditions, including central air conditioning, window/wall units, and evaporative coolers, but excludes non-system fans and blowers. According to Andor et al. (2021) and Wang et al. (2023), measuring household electricity use in physical units (kWh) is widely regarded as more accurate and behaviorally informative than relying on expenditure-based measures. Physical consumption captures actual energy use, independent of price variability, billing structures, and regional subsidies that often distort expenditure data.

Figure 1 presents the distribution of household electricity consumption (kWh), weighted to reflect the national population. The distribution is unimodal and approximately normal, with mild right skewness, indicating that while most households cluster around the center, a smaller proportion consumes substantially higher amounts of electricity. The median consumption is 7.43, with the 25th and 75th percentiles at 6.81 and 7.98, respectively. These correspond to approximately 1,684 kWh, 907 kWh, and 2,947 kWh, highlighting substantial variation in cooling-related electricity usage across U.S. households.

Figure 1: Distribution of U.S. household electricity consumption



3.4 Descriptive Analysis of Cooling Demand

Cooling Degree Days (CDD) serve as the primary climate metric for assessing household cooling needs. CDD measures the cumulative annual temperature deviation above a 65 °F threshold, representing the intensity of cooling requirements across locations. For each household in the 2020 RECS sample, annual CDD values are matched from the nearest National Climate Data Center station, capturing regional variation in climate exposure. Widely applied in energy and climate studies (Cartalis et al. 2001; Sailor & Pavlova 2003; Isaac & Van Vuuren 2009; Petri & Caldeira 2015; Emenekwe & Emodi 2022), CDD provides a consistent proxy for quantifying climate-driven cooling demand.

Figure 3.2 illustrates the distribution of CDD across U.S. households, showing a unimodal, slightly right-skewed pattern with a median of 7.20 (≈ 1,341 CDD) and an interquartile range from 6.86 (954 CDD) to 7.63 (2,062 CDD). Panel B displays a similar but narrower spread, from 6.66 (778 CDD) to 7.55 (1,889 CDD). Table 3 reports summary statistics by climate zone: households in Hot Humid regions record the highest average CDD (3,310), followed by Hot & Mixed Dry (2,254) and Mixed Humid (1,501), while Cold & Very Cold and Marine zones show substantially lower averages (896 and 637, respectively). The 30-year averages follow the same pattern with slightly lower magnitudes, reflecting spatial disparities in cooling demand across U.S. climates.

Figure 2: Distribution of 2020 CDD and 30yrs average CDD among households

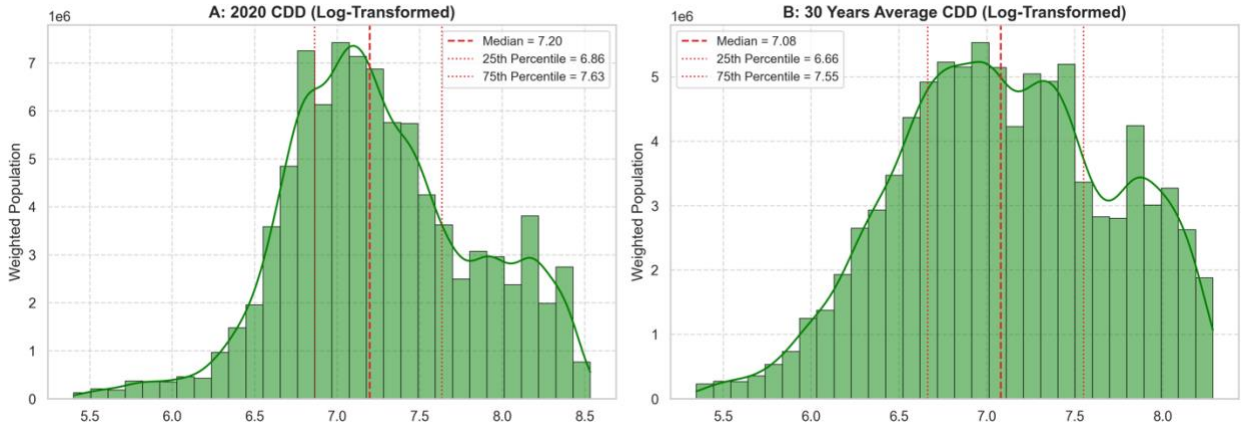


Table 3: Weighted descriptive statistics of cooling degree days by climate zones

	Count	Mean	Std Dev	Min	25th Percentile	Median	75th Percentile	Max
Panel A: 2020 Cooling Degree Days								
Marine	306	636.89	411.31	221.00	331.74	477.73	846.21	2168.00
Cold & Very Cold	5407	896.32	225.35	228.00	758.00	892.00	1046.00	2644.00
Hot Humid	1780	3310.29	689.76	1773.00	2744.09	3313.00	3760.83	4896.00
Hot & Mixed Dry	985	2254.37	976.62	553.00	1643.41	2001.73	2666.00	5082.00
Mixed Humid	4720	1500.78	335.43	404.00	1265.00	1460.00	1720.00	2748.00
Panel B: 30 Years Average Cooling Degree Days								
Marine	306	464.56	216.46	209.00	316.25	399.56	586.67	1741.00
Cold & Very Cold	5407	727.34	223.74	209.00	562.00	722.48	882.00	2276.00
Hot Humid	1780	2858.46	496.88	1591.00	2497.00	2812.00	3214.00	3965.00
Hot & Mixed Dry	985	1752.96	826.53	283.00	1216.05	1545.05	2015.25	3965.00
Mixed Humid	4720	1415.48	373.34	302.00	1128.40	1403.00	1671.00	2715.00

3.5 Control & Moderating Variables

To ensure robust identification of the impact of cooling demand on household electricity consumption, the model includes a set of control and moderating variables. Continuous controls are heating degree days (HDD), representing annual cold-weather exposure; cooled square footage and total number of rooms, reflecting dwelling size and energy demand; household size, capturing occupancy effects; and the number of light bulbs used for one to four hours daily, serving as a proxy for lighting intensity. Categorical controls include air-conditioning type (central, unitary, or evaporative), household income brackets, race (White, Black or African American, American

Indian or Alaska Native, Asian, Pacific Islander, or Multiracial), housing type (single-family, apartment, or mobile home), and insulation presence.

In the subsequent analysis, heterogeneity in the effect of cooling demand is examined across key household characteristics, including income group, air-conditioning type, race, and housing type. These variables provide the basis for assessing subgroup-specific differences in energy responses to rising cooling needs and their selection align with established energy demand literature linking climatic (Sailor & Pavlova 2003; Isaac & Van Vuuren 2009), structural (Bedir et al. 2013; Kavousian et al. 2013), and socioeconomic factors (Hernández 2013; Bednar & Reames 2020; Petri & Caldeira 2015) to household energy use.

Table 5 summarizes household electricity consumption across these covariates. Households in hotter climates, such as the Hot Humid zone, consume significantly more electricity (3,705 kWh) than those in Marine or Cold zones. Electricity use also increases with income, from 1,824 kWh among the lowest earners to 2,426 kWh for households earning over $100k. Single-family homes and central AC users exhibit higher consumption than apartments or users of unitary AC and evaporative coolers. Racial disparities are also evident, with Pacific Islander and White households using more electricity on average than Asian households.

Table 5: Descriptive statistics of household electricity consumption by projected covariates

	Count	Mean	Std Dev	Min	25th Percentile	Median	75th Percentile	Max
Panel A: Climate Zone								
Marine	306	1028.63	865.99	107.91	411.19	779.77	1337.81	7314.97
Cold & Very Cold	5407	1345.91	965.13	76.32	649.66	1102.12	1790.28	8308.31
Hot Humid	1780	3704.98	2043.78	124.91	2116.83	3410.17	5020.31	10124.60
Hot & Mixed Dry	985	2168.34	1727.78	92.65	870.20	1628.69	3054.72	9478.86
Mixed Humid	4720	2220.01	1514.03	113.01	1095.37	1872.65	2987.02	10031.42
Panel B: Income Group								
Less–20k	1492	1824.00	1552.35	76.32	721.72	1362.82	2413.57	9511.89
20k–39k	2333	1980.90	1555.66	102.26	824.67	1527.33	2765.87	9615.44
40k–59k	2105	2146.69	1747.09	90.43	868.98	1633.27	2933.02	9890.57
60k–79k	1458	2164.06	1660.75	110.12	908.98	1766.46	2870.83	8661.61
80k–99k	1758	2158.46	1620.37	98.23	977.75	1701.31	2927.32	10031.42
100k+	4052	2425.62	1842.85	86.63	1075.94	1903.82	3263.88	10124.60
Panel C: Housing Type								
Apartment Home	2235	1237.30	987.97	76.32	569.10	963.61	1636.76	8108.06
Mobile Home	699	2344.84	1745.06	102.69	956.02	1904.85	3231.89	9615.44
Single family Home	10264	2442.24	1777.13	90.43	1118.19	1978.45	3301.39	10124.60
Panel D: Air Conditioner Type								
Central AC	10329	2376.71	1757.04	86.63	1064.15	1898.99	3224.27	10124.60
Unitary AC	2519	1373.55	1208.45	76.32	576.46	1008.11	1782.99	9615.44
Evaporative Coolers	350	1319.34	1105.18	103.87	560.14	1001.13	1648.92	6482.54
Panel E: Household Race								
Black	1083	2207.40	1710.43	92.65	959.03	1734.44	3030.21	9816.55
Multiracial	276	2185.36	1544.20	102.69	970.94	1896.68	2933.50	7589.40
Native American	92	1901.65	1567.53	141.10	842.71	1508.40	2392.86	7390.69
Pacific Islander	21	2540.97	1641.74	512.81	963.77	2165.57	3574.48	6793.60
White	11274	2165.38	1705.94	76.32	915.18	1683.78	2942.36	10124.60
Asian	452	1762.34	1657.02	105.42	628.67	1181.84	2340.36	9329.33

3.6 Empirical Model

To capture the heterogeneous impacts of climate change–induced cooling demand on household electricity consumption, this study employs an econometric strategy grounded in the potential outcome framework. The approach conceptualizes electricity consumption under treated and untreated states, $Y_i(1)$ and $Y_i(0)$, enabling causal analysis of cooling exposure while allowing treatment effects to vary across households and along the electricity consumption distribution.

Identification relies on rich household- and housing-level covariates, normalization of treatment within climate zones, and overlap diagnostics that support comparability between treated and control households.

The empirical strategy proceeds in multiple stages. First, a binary treatment variable (A_i) is constructed using household-level CDD to capture exposure to elevated cooling demand relative to local climatic conditions. CATE are then estimated using a suite of meta-learners (S-, T-, X-, and DR-learners) that flexibly model heterogeneous responses to cooling demand. The framework is subsequently extended to estimate CDTE, specifically, conditional quantile treatment effects (CQTE) and conditional superquantile treatment effects (CSQTE) to examine how cooling exposure reshapes different points and tails of the electricity consumption distribution. Finally, estimated treatment effects are projected onto key household and building characteristics using linear projections to enhance interpretability and to identify subgroups most affected by rising cooling demand.

3.6.1 Treatment Variable of Interest and Identification Strategy

The identification strategy in this study exploits plausibly exogenous variation in climate-driven cooling exposure across households within the five climate zones. Weather realizations that determine CDD are not chosen by households and are orthogonal to short-run household behavioral decisions, conditional on location and observable housing and socioeconomic characteristics. Consistent with the climate–energy literature, including Deschênes and Greenstone (2011); Auffhammer et al. (2017); and Burke et al. (2015), the analysis treats variation in cooling demand as a quasi-random environmental exposure that can be leveraged for causal inference under a selection-on-observables framework.

To capture variation in climate-driven cooling exposure across regions, this study defines a binary treatment indicator representing high cooling demand. A household i in climate zone z is considered treated if its 2020 Cooling Degree Days ($CDD_{i,2020}$) exceed the mean CDD of its respective climate zone ($\bar{C}DD_{z,2020}$). This zone-specific normalization controls for structural climatic differences and enables meaningful within-zone comparisons of cooling-related electricity consumption, consistent with Auffhammer and Mansur (2014) and Hsiang (2016) that exploits relative weather intensity rather than absolute temperature levels. A binary treatment

definition is adopted to facilitate causal interpretation and compatibility with meta-learners and distributional treatment effect estimators, while preserving sufficient within-zone variation for identification. Formally, the treatment indicator is defined as:

$$A_{iz} = \begin{cases} 1, \ if \ CDD_{i,2020} > \bar{C}DD_{z,2020} \\ 0, \ otherwise \end{cases}$$

The identifying assumption underlying this design is that, conditional on observed covariates, households above and below the zone-specific CDD threshold are comparable in terms of underlying determinants of electricity consumption, such that differences in outcomes can be attributed to differential exposure to high cooling demand. This conditional unconfoundedness assumption is standard in observational studies such as Rosenbaum and Rubin (1983); and Imbens and Rubin (2015) that uses weather variation as treatment.

To assess the validity of this quasi-experimental treatment assignment, this study conducts balance checks using pre-treatment covariates. These covariates reflect key determinants of energy consumption and include indicators for climate zones, types of air conditioning equipment, household income ranges, racial identification of the household head, housing unit types, insulation levels, log-transformed cooled square footage, total number of rooms, household size, number of light bulbs used for 1–4 hours daily, and heating degree days (HDD). Balance is assessed by comparing treated and control households—those above and below the zone-specific mean CDD—within symmetric bandwidths of ±75, ±100, and ±125 CDDs. For each specification, two-sample t-tests and standardized mean differences (SMDs) are used to evaluate the similarity of covariates across groups.

SMDs are used as the primary balance metric because they are sample-size invariant and directly quantify covariate imbalance, consistent with best practices in causal inference and and the broader tradition of balance diagnostics in observational designs (Rosenbaum & Rubin 1985; Ho et al. 2007; Austin 2009; Stuart 2010; Imbens & Rubin 2015; Cattaneo & Titiunik 2022). Across all three bandwidths, SMDs remain well within conventional thresholds for acceptable balance, supporting the plausibility of quasi-random exposure to high cooling demand around the climate-zone CDD threshold. Detailed balance statistics are reported in Appendix Table A3.1.

For robustness, two alternative classification strategies are employed to ensure that the results are not driven by a single threshold definition or short-term climatic variation. First, households are reclassified as treated if their 2020 Cooling Degree Days ($CDD_{i,2020}$) exceed the 75th percentile of CDD within their respective climate zone ($CDD^{(75)}_{z,2020}$). This specification isolates households facing especially high cooling exposure relative to their local climatic context and allows the analysis to assess sensitivity to more extreme treatment definitions. Formally, this alternative treatment indicator is defined as:

$$A^{(75)}_{iz} = \begin{cases} 1, \ if \ CDD_{i,2020} > CDD^{(75)}_{z,2020} \\ 0, \ otherwise \end{cases}$$

Second**,** the treatment definition is repeated using long-run climate exposure based on 30-year average Cooling Degree Days ($CDD_{i,30yr}$). This long-term specification captures persistent climatic exposure and validates the stability of the results under extended temperature trends rather than annual anomalies. In this case, a household is treated if its 30-year average CDD exceeds the mean 30-year CDD of its climate zone ($\bar{C}DD_{z,30yr}$).

$$A^{(30)}_{iz} = \begin{cases} 1, \ if \ CDD_{i,30yr} > \bar{C}DD_{z,30yr,} \\ 0, \ otherwise \end{cases}$$

These alternative treatment definitions allow the analysis to distinguish between short-run exposure to elevated cooling demand and long-run climatic intensity, a distinction supported by Dell et al. (2014) and Hsiang et al. (2018). This provides a robustness check on the stability of the estimated treatment effects. Subsequent sections quantify how exposure to high cooling demand causally affects household electricity consumption across the conditional mean and the full outcome distribution using causal machine learning methods.

3.6.2 Conditional average treatment effects specification

The potential outcomes framework defines electricity consumption for each household under two scenarios: $Y_i(1)$ represents the electricity use if household i experiences high cooling energy demand (treated), while $Y_i(0)$ represents the electricity use if household i does not experience high cooling energy demand (control). The CATE (τ) is defined as the expected difference in electricity consumption under these two scenarios, conditional on household characteristics (X_i):

$$CATE(X_i) = E[Y_i(1) - Y_i(0)|X_i] \qquad (1)$$

where, X_i includes household-level covariates such as air conditioning type, household income, cooled sqft, household race, etc. These covariates help to capture the heterogeneity in treatment effects across different subpopulations.

The observed outcome is defined as:

$$Y_i = A_i Y_i(1) + (1 - A_i) Y_i(0) \qquad (2)$$

where A_i is the binary indicator for high cooling demand. Under the standard assumptions of conditional unconfoundedness $(Y(a) \perp A|X)$ and overlap $(0 < P(A = 1|X) < 1)$, CATEs are identified from observational data.

To estimate CATEs, this study employs a suite of meta-learners—the S-learner, T-learner, X-learner, and doubly robust (DR) learner following the framework of Künzel et al. (2019). These learners differ in how they model treated and control potential outcomes and in their robustness to treatment imbalance and model misspecification. These learners provide complementary approaches on heterogeneous responses to cooling demand in a high-dimensional household energy setting.

The S-learner estimates a single outcome model that includes treatment status as a covariate, offering computational efficiency but imposing a common functional form across treatment states. The T-learner estimates separate outcome models for treated and control households, allowing greater flexibility at the cost of higher variance when group sizes are uneven. The X-learner builds on the T-learner by imputing individual-level treatment effects and re-weighting them to improve performance under treatment imbalance. Finally, the DR-learner combines outcome modeling with propensity score adjustment, yielding robust CATE estimates that remain consistent when either the treatment model or outcome model is correctly specified. These learners are particularly well suited to residential electricity demand, where climate exposure interacts nonlinearly with income, housing characteristics, and cooling technology. To avoid overfitting and regularization bias, all learners are implemented with cross-fitting and flexible machine-learning base models. Full algorithmic details for each learner are provided in Appendix A1.

3.6.3 Conditional distributional treatment effects specification

Extending the same household-level treatment indicator A_i defined in Section 3.6.1, the analysis adopts a distributional approach to examine how elevated cooling demand reshapes the entire conditional distribution of household electricity consumption, rather than only its mean. While CATE estimators capture average heterogeneity across observed characteristics, they do not reveal whether treatment effects are concentrated in the lower or upper tails of the outcome distribution. Distributional treatment effects address this limitation by characterizing how climate-driven cooling exposure affects households differently across the full consumption distribution. Building on the framework of Kallus and Oprescu (2023), CDTEs quantify differences in specific distributional statistics such as quantiles and super-quantiles between treated and control potential outcome distributions.

Formally, the CDTE is as:

$$CDTE(X) = \aleph^*\left(F_{Y(1)|X}\right) - \aleph^*\left(F_{Y(0)|X}\right) \qquad (3)$$

where $\aleph^*(F)$ denotes a function of the conditional outcome distribution defined through a moment condition,

$$E[\varrho(Y, \aleph, h)] = 0$$

and $F_{Y(a)|X}$ represents the conditional distribution of electricity consumption under treatment state $a \in \{0,1\}$.

The study employs two key CDTE measures to assess the impact of cooling energy demand on electricity use. The first measure is the conditional quantile treatment effect (CQTE), which captures the treatment effect at specific quantiles of the electricity use distribution. It is defined as:

$$CQTE(X;\tau) = q\left(F_{Y(1)|X};\tau\right) - q\left(F_{Y(0)|X;\tau}\right) \qquad (4)$$

where $q(F;\tau)$ denotes the $\tau - th$ quantile. CQTEs show how high cooling demand differentially affects households at the lower and upper tails of electricity consumption.

The second measure is the conditional super-quantile treatment effect (CSQTE), which characterizes treatment effects in the upper tail by comparing expected outcomes beyond a given quantile:

$$CSQTE(X;\tau) = \mu\left(F_{Y(1)|X};\tau\right) - \mu\left(F_{Y(0)|X;\tau}\right) \qquad (5)$$

Here, $\mu(F;\tau)$ represents the expected value of outcomes exceeding the τ-quantile. CSQTEs are particularly relevant for identifying households facing extreme cooling-related electricity burdens, which are central to concerns about energy affordability, peak-load stress, and climate adaptation.

Estimation of CDTEs follows the pseudo-outcome regression approach proposed by Kallus and Oprescu (2023), which ensures robustness to nuisance parameter estimation and model misspecification through orthogonalization and cross-fitting. The full pseudo-outcome construction and nuisance estimation steps are presented in Appendix A2.

3.6.4 Linear projections of CATE, CQTE, and CSQTE estimates

After computing CATE, CQTE, and CSQTE estimates, this study examines treatment effect heterogeneity by projecting the estimated values from each estimator onto household and regional covariates. The estimated coefficients β_j summarize how treatment effects vary with observable household and housing characteristics. These projections are used as descriptive summaries of heterogeneity rather than as causal regressions and do not introduce additional identifying assumptions beyond those required for the underlying treatment effect estimators. The projection is performed separately for each categorical group to identify the characteristics most strongly associated with variation in treatment effects across estimators. This separate-group projection strategy is adopted to minimize multicollinearity between categorical variables and to enhance interpretability by isolating the contribution of each group to observed heterogeneity.

The empirical projection is defined as:

$$\widehat{\tau_i} = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \varepsilon_i \qquad (6)$$

where $\widehat{\tau_i}$ denote a fitted, unit-specific treatment-effect estimate from any of the three estimators (CATE, CQTE at a fixed q, or CSQTE at a fixed q), X_{ij} are indicator variables drawn from one categorical group at a time (household air-conditioning types, race, income group, or household building type)**,** and ε_i is the error term.

Equation (19) serves as the sample analogue of the population best linear projection established in the CDTE framework:

$$\beta^* = \underset{\beta \epsilon Rp}{arg\ min} \,||\, CDTE - \phi(X)^T \beta \,||\qquad (7)$$

where $\tau(X)$ represents the true conditional effect (CATE, $CQTE_q$, or $CSQTE_q$) and $\phi(X)$ denotes the projected covariates. Under regularity and cross-fitting conditions, the coefficient vector $\hat{\beta}$ converges to β^* and satisfies:

$$\sqrt{n}\left(\hat{\beta} - \beta^*\right) \xrightarrow{d} N(0, \Sigma) \qquad (8)$$

where Σ is the asymptotic covariance matrix of the regression of the pseudo-outcome on $\phi(X)$. HC3-robust standard errors are reported in empirical applications. Since the dependent variable in the original model is in logarithmic form, coefficient magnitudes can be interpreted as semi-elasticities, and percentage effects can be obtained via $(exp(\hat{\tau}) - 1)100$ if desired.

3.6.5 Estimation approach

The estimation of heterogeneous average and distributional treatment effects follows recent advances in causal machine learning that combine pseudo-outcome regression, orthogonalization, and cross-fitting to support valid causal inference in high-dimensional observational settings. The analysis jointly estimates CATE and CDTE, allowing treatment effects of high cooling demand to vary flexibly across households and across the electricity consumption distribution. These estimators are implemented within a cross-fitted framework to maintain orthogonality between nuisance estimation and treatment effect.

Consistent with practices in causal machine learning from Chernozhukov et al. (2018) and Athey and Imbens (2019), to implement cross-fitting, the full analytical sample is partitioned into

mutually exclusive training and estimation subsamples. Nuisance components—including the propensity score and outcome regression functions—are trained on auxiliary subsamples, while treatment effects are estimated on held-out data. This procedure ensures that the same observations are not used simultaneously for model training and effect estimation, thereby reducing overfitting bias and producing orthogonalized treatment effect estimates.

Three complementary objects are estimated, CATE, CQTE, and CSQTE. CATEs are estimated using multiple meta-learners, including the S-, T-, X-, and doubly robust (DR) learners to flexibly model heterogeneous responses to cooling demand. Employing multiple learners reduces dependence on any single modeling assumption and provides a robustness check on estimated heterogeneity patterns. Distributional effects (CQTE and CSQTE) are estimated using the model-agnostic CDTE framework, which enables valid inference without imposing restrictive functional form assumptions. In a final interpretive step, estimated CATE, CQTE, and CSQTE values are projected onto selected household and housing characteristics using OLS. All estimation procedures are implemented in Python, drawing on the meta-learning framework of Künzel et al. (2019) and the CDTE learner of Kallus and Oprescu (2023). Additional computational and software implementation details are provided in Appendix A3.

4.0 Empirical Findings

Figure 3 presents the distribution of estimated treatment effects across the full sample, combining results from four CATE learners (S, T, X, and DR) and the distributional estimators (CQTE and CSQTE) evaluated at the 25th and 75th percentiles. Across the CATE estimators, the distributions are tightly centered and highly consistent, with median effects clustered around 0.28–0.29. This stability across learners indicates a robust positive average response to elevated cooling demand. At the median, this corresponds to an increase in electricity consumption of approximately 32–34 percent for households exposed to higher-than-average cooling demand relative to comparable households within the same climate zone.

Moving beyond average effects, the CQTE and CSQTE distributions show meaningful heterogeneity across the consumption distribution. The CQTE estimates show larger effects at the lower quantile than at the upper quantile, indicating that proportional responses to increased cooling demand are stronger among lower-consuming households. The CSQTE estimates exhibit

a similar pattern but with greater dispersion, reflecting heightened variability in tail outcomes. At the upper tail (75th percentile), the median CSQTE is approximately 0.14, implying an increase in electricity consumption of about 15 percent among households in the high-consumption tail. Notably, the upper-tail CSQTE distributions display longer right tails, highlighting the presence of a subset of households for whom cooling demand translates into particularly large increases in electricity use.

Figure 3: Distribution of estimated effects under increased cooling demand

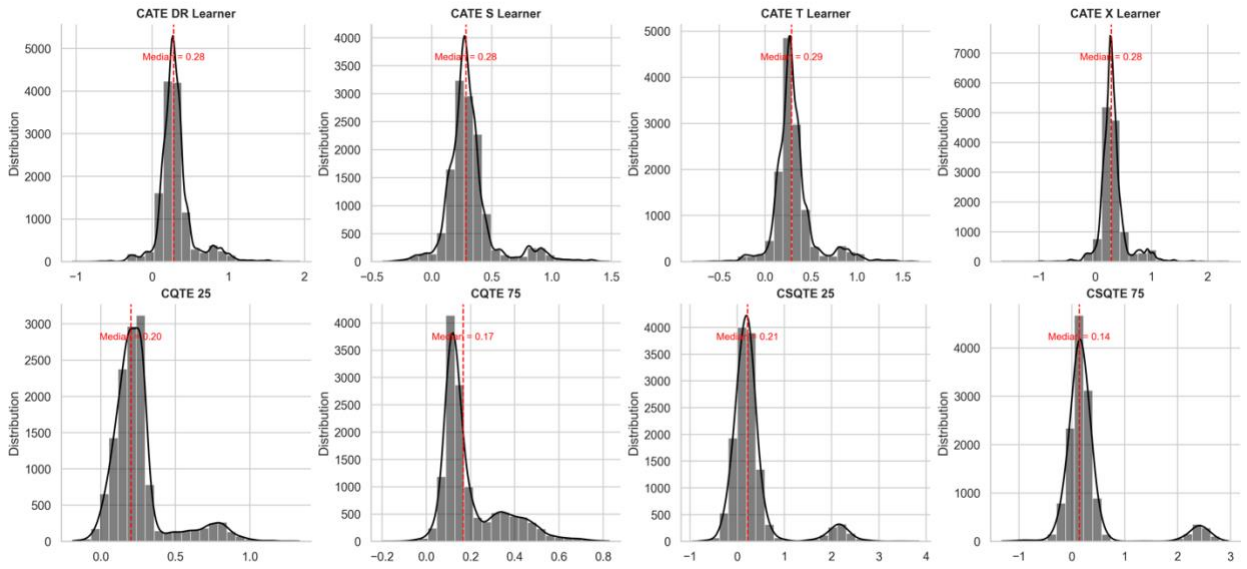


Table 6 reports OLS projections of the estimated CATE, CQTE, and CSQTE onto air-conditioning system types, with unitary AC serving as the reference category. Across all CATE learners, the coefficients for both central AC and evaporative cooler households are positive and statistically significant, indicating that households with more cooling-intensive technologies experience systematically larger treatment effects from elevated cooling demand. The consistency of signs and relative ordering across learners suggests that the observed patterns are not driven by a particular estimator. At the mean, evaporative cooler households exhibit the strongest average response, followed by central AC households, confirming that cooling technology is a key mediator of climate-driven electricity demand.

The distributional projections show important differences between CQTE and CSQTE estimates. In the CQTE columns, central AC households display stronger responsiveness at both lower and upper quantiles, indicating relatively larger effects across the conditional consumption

distribution. In contrast, the CSQTE results show that evaporative cooler households load more heavily in the upper tail, with substantially larger effects among high-consuming households. This divergence highlights that while central AC users respond more uniformly across the distribution, evaporative cooler households face disproportionately larger increases in extreme consumption outcomes. The stability of these patterns across estimators underscores the role of cooling technology in shaping both average and tail vulnerability to rising cooling demand.

Table 6: OLS projection of CATE, CQTE and CSQTE on air condition type

	CATE (S Learner)	**CATE (T Learner)**	**CATE (X Learner)**	**CATE (DR Learner)**	**CQTE (25)**	**CQTE (75)**	**CSQTE (25)**	**CSQTE (75)**
Central AC Type	0.036*** (0.005)	0.041*** (0.006)	0.042*** (0.005)	0.044*** (0.006)	0.048*** (0.011)	-0.087*** (0.012)	0.113*** (0.003)	-0.067*** (0.003)
Evaporative coolers	0.106*** (0.012)	0.126*** (0.015)	0.121*** (0.014)	0.154*** (0.019)	-0.061** (0.040)	-0.115** (0.049)	0.068*** (0.008)	0.074*** (0.009)
Constant	0.283*** (0.004)	0.279*** (0.005)	0.278*** (0.005)	0.275*** (0.006)	0.289*** (0.010)	0.382*** (0.010)	0.162*** (0.003)	0.246*** (0.003)
R^2	0.009	0.008	0.010	0.011	0.002	0.003	0.053	0.057
N	13198	13198	13198	13198	13198	13198	13198	13198
Standard errors in parentheses. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.								

Table 7 presents OLS projections of the estimated CATE, CQTE, and CSQTE onto household income groups, with households earning below $20,000 serving as the reference category. Across all three estimators, the projected coefficients are positive for every income group, indicating that elevated cooling demand increases electricity consumption regardless of income. However, the magnitude of these effects varies non-monotonically across income levels. In the CATE projections, middle-income households, particularly those in the $60,000–$79,000 range exhibit the strongest average response relative to the lowest-income group, while upper-middle income households show comparatively smaller effects before responsiveness increases again among the highest-income households. This pattern suggests that sensitivity to cooling demand is not strictly increasing in income and is most pronounced among middle-income households.

The distributional projections show further nuance in how these income gradients manifest across the consumption distribution. In the CQTE, middle-income households respond strongly at both lower and upper quantiles, indicating larger conditional effects throughout the distribution, while lower-income households exhibit comparatively muted responses, particularly at higher quantiles. In contrast, the CSQTE estimates display a flatter profile across income groups, suggesting that

tail-averaged effects are less sharply differentiated by income. These findings show that income-based heterogeneity in cooling demand responses is most pronounced in average and quantile-specific effects rather than in extreme tail outcomes.

Table 7: OLS projection of CATE, CQTE and CSQTE on household income

	CATE (S Learner)	**CATE (T Learner)**	**CATE (X Learner)**	**CATE (DR Learner)**	**CQTE (25)**	**CQTE (75)**	**CSQTE (25)**	**CSQTE (75)**
20k-39k income range	0.043*** (0.006)	0.040*** (0.007)	0.040*** (0.007)	0.046*** (0.007)	0.195*** (0.019)	0.100*** (0.021)	0.060*** (0.006)	0.013** (0.005)
40k-59k income range	0.027*** (0.007)	0.023*** (0.007)	0.023*** (0.007)	0.021*** (0.008)	0.385*** (0.019)	0.176*** (0.021)	0.036*** (0.007)	0.036*** (0.005)
60k-79k income range	0.057*** (0.007)	0.051*** (0.008)	0.050*** (0.008)	0.051*** (0.009)	0.206*** (0.021)	0.257*** (0.023)	0.069*** (0.007)	0.019*** (0.005)
80k-99k income range	-0.066*** (0.007)	-0.073*** (0.008)	-0.074*** (0.007)	-0.069*** (0.008)	0.187*** (0.019)	0.261*** (0.021)	0.040*** (0.007)	0.002 (0.005)
100k and above income range	0.011 (0.006)	0.003 (0.006)	0.001 (0.006)	0.011 (0.007)	0.181*** (0.017)	0.072*** (0.018)	0.061*** (0.006)	0.011*** (0.004)
Constant	0.301*** (0.005)	0.307*** (0.005)	0.308*** (0.005)	0.303*** (0.006)	0.126*** (0.014)	0.180*** (0.016)	0.204*** (0.005)	0.182*** (0.004)
R^2	0.029	0.024	0.027	0.022	0.031	0.019	0.011	0.006
N	13198	13198	13198	13198	13198	13198	13198	13198

Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

When projecting estimated effects onto household race (Table 8), using Black households as the reference group, the results show consistent but uneven racial heterogeneity in cooling-demand responsiveness across CATE, CQTE, and CSQTE estimators. In the CATE projections, Asian households exhibit effects comparable to Black households, indicating similar average sensitivity to elevated cooling demand, while White households show systematically smaller average effects, suggesting lower responsiveness. This ranking is stable across learners. The distributional estimates show sharp contrasts: CQTE effects are notably stronger for Asian households at lower quantiles, indicating heightened sensitivity among lower-consuming Asian households, whereas White households exhibit weaker responses across both lower and upper quantiles.

The CSQTE projections further highlight racial disparities in tail outcomes. Relative to Black households, Multiracial households show limited differences in average effects but respond more strongly in the lower tail, while Native American and Pacific Islander households exhibit consistently larger tail-averaged effects, indicating heightened vulnerability at extreme consumption levels. Across race groups, heterogeneity is more pronounced in CQTE and CSQTE

than in CATE, underscoring that racial disparities in cooling-related electricity burden are concentrated in the distributional tails rather than in average effects alone. The consistency of these patterns across estimators strengthens the evidence that climate-driven cooling demand disproportionately affects certain minority households, particularly in high-risk consumption households.

Table 8: OLS projection result of CATE, CQTE and CSQTE on household race

	CATE (S Learner)	**CATE (T Learner)**	**CATE (X Learner)**	**CATE (DR Learner)**	**CQTE (25)**	**CQTE (75)**	**CSQTE (25)**	**CSQTE (75)**
Asian Households	0.009 (0.020)	0.010 (0.024)	-0.007 (0.022)	-0.007 (0.025)	0.092*** (0.040)	0.035*** (0.043)	0.062*** (0.015)	0.051*** (0.009)
Multiracial Households	0.027 (0.021)	0.005 (0.025)	0.010 (0.024)	-0.008 (0.027)	0.053 (0.049)	0.192*** (0.056)	0.091*** (0.015)	0.048*** (0.010)
Native American Households	0.122*** (0.033)	0.079 (0.049)	0.089** (0.044)	0.044 (0.063)	-0.497*** (0.080)	-0.886*** (0.085)	0.107*** (0.022)	0.094*** (0.018)
Pacific Islander Households	0.136** (0.099)	0.180*** (0.120)	0.104** (0.122)	-0.004 (0.150)	0.155*** (0.202)	0.165* (0.235)	0.205*** (0.067)	0.062 (0.043)
White Households	-0.026*** (0.008)	-0.045*** (0.009)	-0.044*** (0.008)	-0.041*** (0.009)	-0.117*** (0.017)	0.006 (0.019)	-0.012 (0.007)	0.021*** (0.004)
Constant	0.334*** (0.007)	0.351*** (0.009)	0.351*** (0.008)	0.349*** (0.009)	0.407*** (0.016)	0.293*** (0.018)	0.257*** (0.006)	0.174*** (0.004)
R^2	0.009	0.012	0.011	0.004	0.063	0.032	0.015	0.006
N	13198	13198	13198	13198	13198	13198	13198	13198

Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 9 presents OLS projections of the estimated CATE, CQTE, and CSQTE onto home types, with apartment homes serving as the reference category. The result show clear heterogeneity in cooling-demand responsiveness across housing types, with patterns that differ across CATE, CQTE, and CSQTE estimators. In the CATE projections, households residing in mobile homes and single-family homes both exhibit larger average increases in electricity consumption relative to apartment homes, indicating greater exposure to cooling-related energy demand. Across learners, the sign and ranking of these effects are consistent, with mobile homes showing the largest average response and single-family homes also responding strongly relative to apartments.

The distributional estimates show important contrasts in where these housing effects concentrate along the consumption distribution. In the CQTE projections, single-family homes show strong and statistically significant effects at both lower and upper quantiles, suggesting heightened

responsiveness across the distribution. In contrast, mobile homes exhibit weaker and generally insignificant CQTE effects, indicating limited sensitivity at specific quantiles despite larger average effects. The CSQTE results reinforce this distinction: single-family homes respond more strongly in the tails, while mobile-home effects attenuate, particularly in the upper tail. These results suggest that single-family households face disproportionate cooling burdens across both average and extreme consumption household, whereas mobile-home vulnerability appears concentrated in mean effects rather than tail outcomes.

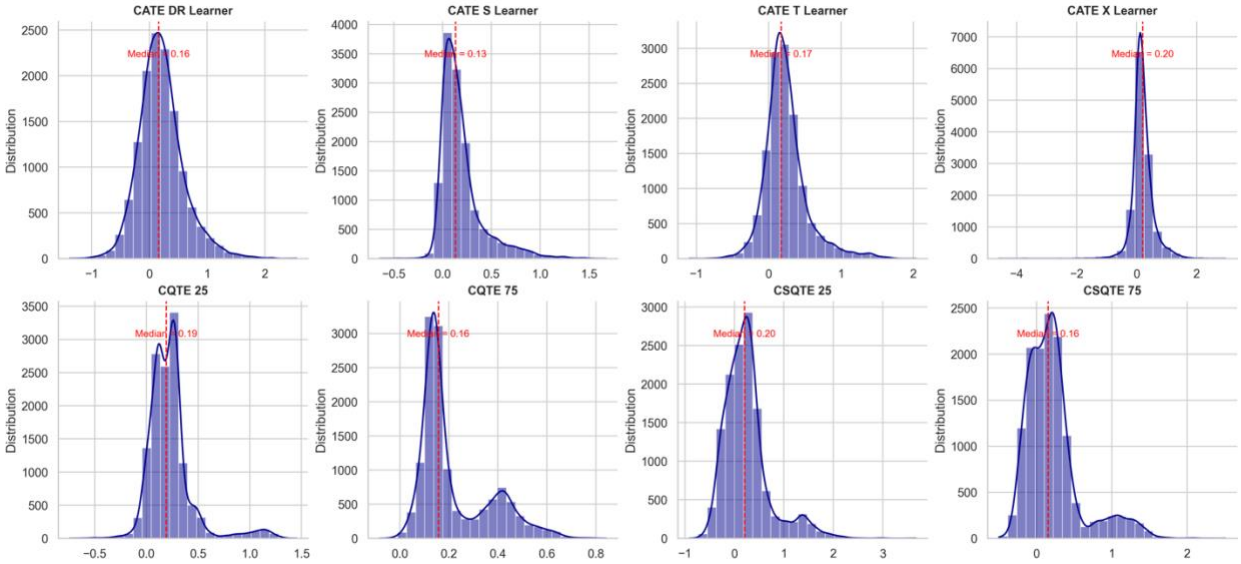
Table 9: OLS projection of CATE, CQTE and CSQTE on building types

	CATE (S Learner)	**CATE (T Learner)**	**CATE (X Learner)**	**CATE (DR Learner)**	**CQTE (25)**	**CQTE (75)**	**CSQTE (25)**	**CSQTE (75)**
Mobile Homes	0.121*** (0.009)	0.141*** (0.011)	0.136*** (0.010)	0.131*** (0.012)	0.023 (0.025)	0.009 (0.028)	0.018** (0.008)	-0.001 (0.007)
Single Family Homes	0.085*** (0.006)	0.091*** (0.007)	0.083*** (0.006)	0.084*** (0.007)	0.211*** (0.014)	0.205*** (0.015)	0.093*** (0.005)	0.012*** (0.003)
Constant	0.241*** (0.006)	0.236*** (0.006)	0.243*** (0.006)	0.241*** (0.007)	0.159*** (0.013)	0.152*** (0.013)	0.179*** (0.005)	0.187*** (0.003)
R^2	0.026	0.025	0.024	0.020	0.023	0.018	0.037	0.001
N	13198	13198	13198	13198	13198	13198	13198	13198

Standard errors in parentheses. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Figure 4 examines the robustness of the baseline results by redefining treatment as exposure to extreme cooling demand, defined as household-level 2020 cooling degree days exceeding the 75th percentile within each climate zone. Under this stricter treatment definition, the distributions of CATE estimates remain positive and tightly clustered across learners, confirming that the direction and presence of treatment effects are not sensitive to the choice of threshold. CQTE estimates continue to exhibit clear heterogeneity across the outcome distribution, with effects systematically larger at lower quantiles than at upper quantiles. CSQTE estimates display wider dispersion, reflecting greater variability in tail outcomes under extreme heat exposure. Overall, the persistence of positive and ordered effects indicates that elevated cooling demand continues to impose disproportionate burdens across the electricity consumption distribution even under more stringent exposure criteria. Additional robustness checks based on long-run (30-year) average cooling degree days yield similar patterns and are reported in Appendix Figure A1.

Figure 4: Robustness analysis of treatment effects under extreme cooling demand



Under the alternative treatment definition based on the 75th percentile of CDD within climate zones, the OLS projections in Table 3.9 confirm the stability of the baseline results across CATE, CQTE, and CSQTE estimators. Across panels, coefficient signs and relative magnitude remain consistent, indicating that the main heterogeneity patterns are not sensitive to how high cooling exposure is defined. For air-conditioning systems (Panel A), central AC households continue to exhibit stronger average and lower-tail responses than the unitary AC baseline, while evaporative cooler households respond more heavily in the upper tail, a pattern that is evident in both CQTE and CSQTE projections.

Panels B–D further demonstrate that the core sources of heterogeneity persist across income, race, and housing type. The non-monotonic income gradient observed in the baseline projections remained, with middle-income households exhibiting stronger responsiveness than both lower- and higher-income groups across estimators. Racial heterogeneity remains pronounced in the tails, with minority households showing systematically larger CQTE and CSQTE effects than White households. Housing-type asymmetries also remain, with single-family homes responding more strongly in lower-tail outcomes and mobile homes showing comparatively stronger upper-tail responses. Overall, these findings confirm that the distributional structure of cooling-demand impacts is stable and not an artifact of the baseline treatment threshold. Additional robustness

check using long-run (30-year) average CDD show similar results and is reported in Appendix Table A2.

Table 10: Heterogeneous effects under 2020 CDD 75th percentile treatment threshold

	CATE (S Learner)	CATE (T Learner)	CATE (X Learner)	CATE (DR Learner)	CQTE (25)	CQTE (75)	CSQTE (25)	CSQTE (75)
Panel A: Air Conditioner Type								
Central AC Type	-0.020*** (0.005)	-0.014 (0.010)	0.052*** (0.007)	0.041*** (0.009)	-0.042*** (0.010)	-0.120*** (0.008)	0.086*** (0.004)	-0.089*** (0.004)
Evaporative coolers	0.047*** (0.014)	0.059** (0.026)	0.048** (0.019)	0.053** (0.023)	-0.138*** (0.026)	0.024 (0.022)	0.068*** (0.010)	0.037*** (0.010)
Constant	0.205*** (0.005)	0.216*** (0.010)	0.198*** (0.007)	0.159*** (0.008)	0.292*** (0.009)	0.314*** (0.007)	0.158*** (0.003)	0.289*** (0.004)
R^2	0.003	0.002	0.005	0.002	0.002	0.017	0.023	0.074
N	13198	13198	13198	13198	13198	13198	13198	13198
Panel B: Income Group								
$20k–$39k	0.040*** (0.007)	0.073*** (0.013)	0.060*** (0.010)	0.064*** (0.012)	0.116*** (0.015)	0.072*** (0.012)	0.053*** (0.007)	0.020*** (0.005)
$40k–$59k	0.032*** (0.007)	0.006 (0.013)	0.012 (0.010)	0.027** (0.012)	0.105*** (0.015)	0.100*** (0.012)	0.030*** (0.007)	0.017*** (0.005)
$60k–$79k	0.070*** (0.008)	0.092*** (0.015)	0.088*** (0.011)	0.081*** (0.013)	0.321*** (0.017)	0.295*** (0.013)	0.087*** (0.008)	0.040*** (0.005)
$80k–$99k	0.040*** (0.007)	-0.002 (0.014)	0.006 (0.010)	0.010 (0.012)	0.020 (0.016)	0.002 (0.013)	0.029*** (0.008)	0.005 (0.005)
$100k+	0.040*** (0.006)	0.034*** (0.012)	0.030*** (0.009)	0.037*** (0.011)	0.062*** (0.014)	0.057*** (0.011)	0.027*** (0.007)	-0.006 (0.004)
Constant	0.153*** (0.005)	0.173*** (0.010)	0.208*** (0.008)	0.155*** (0.009)	0.161*** (0.012)	0.142*** (0.009)	0.191*** (0.005)	0.210*** (0.004)
R^2	0.006	0.007	0.009	0.005	0.034	0.046	0.011	0.010
N	13198	13198	13198	13198	13198	13198	13198	13198
Panel C: Household Race								
Asian	0.056*** (0.016)	0.153*** (0.029)	0.086*** (0.022)	0.093*** (0.025)	0.373*** (0.035)	0.279*** (0.028)	0.017 (0.017)	0.060*** (0.009)
Multiracial	0.082*** (0.018)	0.118*** (0.031)	0.124*** (0.024)	0.112*** (0.028)	0.002 (0.035)	-0.159*** (0.030)	0.082*** (0.019)	0.052*** (0.010)
Native American	0.083*** (0.027)	0.166*** (0.048)	0.154*** (0.038)	0.069 (0.038)	0.001 (0.060)	0.216*** (0.052)	0.119*** (0.027)	0.098*** (0.016)
Pacific Islander	0.151 (0.085)	0.117 (0.113)	0.160 (0.100)	0.223 (0.129)	0.113*** (0.142)	0.235*** (0.127)	0.250*** (0.091)	0.072* (0.039)
White	0.037*** (0.006)	0.044*** (0.013)	0.041*** (0.009)	0.037*** (0.011)	0.049*** (0.013)	0.073*** (0.011)	0.038*** (0.007)	0.042*** (0.004)
Constant	0.155*** (0.006)	0.160*** (0.012)	0.198*** (0.008)	0.154*** (0.010)	0.197*** (0.013)	0.150*** (0.011)	0.191*** (0.007)	0.180*** (0.003)
R^2	0.004	0.005	0.005	0.003	0.046	0.029	0.006	0.009
N	13198	13198	13198	13198	13198	13198	13198	13198
Panel D: Housing Type								
Mobile Home	0.055*** (0.008)	0.195*** (0.017)	0.150*** (0.012)	0.111*** (0.015)	-0.145*** (0.020)	0.032* (0.017)	0.075*** (0.009)	0.024*** (0.007)
Single family Home	0.090*** (0.005)	0.107*** (0.010)	0.086*** (0.008)	0.090*** (0.009)	0.085*** (0.012)	-0.009 (0.010)	0.079*** (0.006)	0.000 (0.004)
Constant	0.118*** (0.004)	0.114*** (0.009)	0.165*** (0.007)	0.116*** (0.008)	0.197*** (0.011)	0.226*** (0.009)	0.162*** (0.005)	0.219*** (0.003)
R^2	0.023	0.014	0.015	0.010	0.014	0.001	0.018	0.001

N	13198	13198	13198	13198	13198	13198	13198	13198
Standard errors in parentheses. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.								

In summary, three core empirical insights emerge from the results in this study. First, elevated cooling demand consistently increases household electricity consumption across all estimators, but the magnitude of this effect varies substantially across the consumption distribution. Average effects alone understate this variation: lower-consuming households exhibit stronger proportional responses, while tail-based measures show concentrated burdens among high-use households. Second, heterogeneity is systematic rather than idiosyncratic. Cooling technology, income, race, and housing type repeatedly shape both mean and distributional responses, with the strongest differentiation appearing in CQTE and CSQTE estimates rather than in CATE alone. Third, robustness checks based on alternative high cooling demand definitions and long-run climate exposure confirm that these patterns are stable and structural.

5.0 Conclusion

This study examines the causal and distributional effects of climate-driven cooling demand on household electricity consumption using 2020 Residential Energy Consumption Survey microdata and a causal machine learning framework grounded in the potential outcomes approach. By combining conditional average treatment effects with distributional treatment effect estimators, and subsequently projecting these estimated effects onto observable household characteristics for interpretation, the analysis moves beyond mean responses to characterize how rising cooling demand reshapes the entire distribution of household electricity use within climate zones. This approach enables a more precise assessment of which households bear the greatest marginal and tail-level burdens from increased heat exposure.

The results show that elevated cooling demand causally increases household electricity consumption, with effects that are heterogeneous. While average effects indicate substantial increases in electricity use among households facing higher cooling needs, distributional estimates show that these impacts are unevenly distributed across the consumption spectrum. Lower-consuming households often experience larger proportional increases, while tail-based measures highlight concentrated burdens among high-use households. These findings demonstrate that mean effects alone obscure important dimensions of climate-induced energy vulnerability.

Heterogeneity analyses further show that responses to cooling demand vary systematically across household characteristics. Differences by cooling technology, income, race, and housing structure indicate that climate-related electricity burdens are shaped by both socioeconomic and technological factors. The use of OLS projections of causal machine learning estimates allows these subgroup patterns to be estimated in an interpretable manner, linking causal heterogeneity to concrete household attributes rather than abstract learner outputs. In particular, households with certain cooling systems, middle-income households, and several minority racial groups consistently exhibit stronger average or tail responses to increased cooling demand. These patterns persist across alternative treatment definitions, underscoring that the observed heterogeneity reflects structural differences rather than sensitivity to modeling choices.

Robustness checks confirm the stability of the main findings when cooling exposure is defined using alternative thresholds and long-run climate measures. Across specifications, the direction and relative magnitude of the effects remain consistent, indicating that the estimated heterogeneity is not driven by short-term temperature fluctuations or a particular treatment definition. This stability strengthens the interpretation of the results as reflecting persistent climate-related pressures on household electricity consumption rather than transitory weather shocks.

Overall, this study demonstrates that climate-driven cooling demand imposes substantial and uneven electricity burdens across U.S. households, with important implications for energy affordability and climate adaptation policy. By integrating causal and distributional machine learning methods, the analysis provides a flexible and policy-relevant framework for identifying households most exposed to rising cooling needs. These findings highlight the limits of uniform energy policies and underscore the importance of targeted interventions such as technology-specific efficiency programs and equity-focused adaptation strategies to mitigate the unequal impacts of increasing heat exposure.

References

Auffhammer, Maximilian, Solomon Hsiang, Wolfram Schlenker, and Adam Sobel. 2013. “Using Weather Data and Climate Model Output in Economic Analyses of Climate Change.” *Review of Environmental Economics and Policy* 7 (2): 181–98.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. Generalized random forests. *The Annals of Statistics, 47*(2), 1148–1178. https://doi.org/10.1214/18-AOS1709

Athey, Susan, and Stefan Wager. 2021. "Policy learning with observational data." *Econometrica,* 89(1): 133-161.

Austin, Peter. 2009. "Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples." *Statistics in medicine,* 28(25): 3083-3107.

Bach, Philipp, Victor Chernozhukov, Malte Kurz, and Martin Spindler. 2022. "DoubleML-an object-oriented implementation of double machine learning in python." *Journal of Machine Learning Research,* 23(53): 1-6.

Bednar, Dominic, and Tony Reames. 2020. "Recognition of and response to energy poverty in the United States." *Nature Energy,* 5(6): 432-439.

Belloni, Alexandre, Victor Chernozhukov, Ivan Fernandez-Val, and Christian Hansen. 2017. "Program evaluation and causal inference with high-dimensional data." *Econometrica,* 85(1): 233-298.

Bilello, Stanley Bull, James Ekmann, William Horak, Y. Joe Huang et al. 2008. "Effects of climate change on energy production and use in the United States." *US Department of Energy Publications*, 12.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015. **“**Global Non-Linear Effect of Temperature on Economic Production.” *Nature* 527 (7577): 235–39.

Byambadalai, Undral, Tatsushi Oka, and Shota Yasui. 2024. "Estimating distributional treatment effects in randomized experiments: machine learning for variance reduction." *arXiv preprint arXiv:2407.16037*.

Callaway, Brantly, and Pedro HC Sant’Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of econometrics,* 225(2): 200-230.

Carley, Sanya, and David Konisky. 2020. "The justice and equity implications of the clean energy transition." *Nature Energy,* 5(8): 569-577.

Cattaneo, Matias, and Rocio Titiunik. 2021. "Regression discontinuity designs." *Annual Review of Economics* 14(1): 821-851.

Chen, Jau-er, Chien-Hsun Huang, and Jia-Jyun Tien. 2021. "Debiased/double machine learning for instrumental variable quantile regressions." *Econometrics,* 9(2): 15.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. "Inference on counterfactual distributions." *Econometrica,* 81(6): 2205-2268.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Iván Fernández-Val. 2017. "Fisher-Schultz Lecture: Generic machine learning inference on heterogenous treatment effects in randomized experiments, with an application to immunization in India." *arXiv preprint arXiv:1712.04802*.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/debiased machine learning for treatment and structural parameters." C1-C68.

Chernozhukov, Victor, Ivan Fernandez-Val, and Martin Weidner. 2024. "Network and panel quantile effects via distribution regression." *Journal of Econometrics,* 240(2): 105009.

Curth, Alicia, Richard W. Peck, Eoin McKinney, James Weatherall, and Mihaela van Der Schaar. 2024. "Using machine learning to individualize treatment effect estimation: challenges and opportunities." *Clinical Pharmacology & Therapeutics,* 115(4): 710-719.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. “What Do We Learn from the Weather? The New Climate–Economy Literature.” *Journal of Economic Literature* 52(3): 740–98.

Firpo, Sergio, Nicole Fortin, and Thomas Lemieux. 2009. "Unconditional quantile regressions." *Econometrica,* 77(3): 953-973.

Foster, Dylan, and Vasilis Syrgkanis. 2023. "Orthogonal statistical learning." *The Annals of Statistics,* 51(3): 879-908.

Gadea, Maria Dolores, and Jesus Gonzalo. 2023. "Climate change heterogeneity: A new quantitative approach." *arXiv preprint arXiv:2301.02648*.

Gbadegesin, Tosin, & Nadege Yameogo. 2024. Evaluating the Economic Impacts of the G20 Compact Initiative: Evidence from Causal Inference Using Advanced Machine Learning Techniques. *Journal of Sustainable Development, 17*(5), 1-56.

Gbadegesin, Tosin. 2025. *Impact of Climate-Driven Cooling Demand on U.S. Household Electricity Consumption and Expenditure.* Unpublished manuscript, Howard University.

Giannarakis, Georgios, Vasileios Sitokonstantinou, Roxanne Suzette Lorilla, and Charalampos Kontoes. 2022. "Personalizing sustainable agriculture with causal machine learning." *arXiv preprint arXiv:2211.03179*.

Hahn, Richard, Jared Murray, and Carlos Carvalho. 2020. "Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion)." *Bayesian Analysis,* 15(3): 965-1056.

Hernández, Diana. 2013. "Energy insecurity: a framework for understanding energy, the built environment, and health among vulnerable populations in the context of climate change." *American Journal of Public Health,* 103(4): e32-e34.

Hill, Jennifer. 2011. "Bayesian nonparametric modeling for causal inference." *Journal of Computational and Graphical Statistics,* 20(1): 217-240.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth Stuart. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political Analysis,* 15(3): 199-236.

Hsiang, Solomon M. 2016. “Climate Econometrics.” *Annual Review of Resource Economics* 8(1): 43–75.

Hsiang, Solomon, Robert Kopp, and Amir Jina. 2018. “The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence from 6,700 Cyclones.” National Bureau of Economic Research. No. w20352.

Hsu, Yu-Chin, Martin Huber, and Yu-Min Yen. 2023. "Doubly Robust Estimation of Direct and Indirect Quantile Treatment Effects with Machine Learning." *arXiv preprint arXiv:2307.01049*.

Ichimura, Hidehiko, and Whitney K. Newey. 2022. "The influence function of semiparametric estimators." *Quantitative Economics,* 13(1): 29-61.

Imbens, Guido, and Donald Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jacobsen, Grant. 2019. "Who wins in an energy boom? Evidence from wage rates and housing." *Economic Inquiry,* 57(1): 9-32.

Kallus, Nathan, and Miruna Oprescu. 2023. "Robust and agnostic learning of conditional distributional treatment effects." In *International Conference on Artificial Intelligence and Statistics*, pp. 6037-6060.

Kallus, Nathan, Xiaojie Mao, and Masatoshi Uehara. 2019. "Localized debiased machine learning: Efficient inference on quantile treatment effects and beyond." *arXiv preprint arXiv:1912.12945*.

Kallus, Nathan, Xiaojie Mao, and Masatoshi Uehara. 2024. "Localized debiased machine learning: Efficient inference on quantile treatment effects and beyond." *Journal of Machine Learning Research,* 25(16): 1-59.

Kato, Masahiro, and Masaaki Imaizumi. 2023. "CATE Lasso: conditional average treatment effect estimation with high-dimensional linear regression." *arXiv preprint arXiv:2310.16819*.
Kennedy, Edward. 2023. "Towards optimal doubly robust estimation of heterogeneous causal effects." *Electronic Journal of Statistics,* 17(2): 3008-3049.

Kim, Jee-Seon, Xiangyi Liao, and Wen Wei Loh. 2023. "Assessing cross-level interactions in clustered data using cate estimation methods." In *The Annual Meeting of the Psychometric Society*, pp. 87-97. Cham: Springer Nature Switzerland.

Knaus, Michael. 2021. "A double machine learning approach to estimate the effects of musical practice on student's skills." *Journal of the Royal Statistical Society Series A: Statistics in Society,* 184(1): 282-300.

Knittel, Christopher, and Samuel Stolper. 2025. "Using machine learning to target treatment: The case of household energy use." *The Economic Journal*, ueaf028.

Knittel, Christopher, and Samuel Stolper. 2021. "Machine Learning about Treatment Effect Heterogeneity: The Case of Household Energy Use." *AEA Papers and Proceedings*, 111: 440–44.

Künzel, Sören, Jasjeet Sekhon, Peter Bickel, and Bin Yu. 2019. "Metalearners for estimating heterogeneous treatment effects using machine learning." *Proceedings of the National Academy of Sciences,* 116(10): 4156-4165.

Li, Danny, Liu Yang, and Joseph Lam. 2013. "Zero energy buildings and sustainable development implications–A review." *Energy,* 54: 1-10.

Linden, Ariel, and Paul Yarnold. 2018. "Combining machine learning and propensity score weighting to estimate causal effects in multivalued treatments." *Journal of Evaluation in Clinical Practice,* 22(6): 875-885.

Machluf, Oshri, Tzviel Frostig, Gal Shoham, Tomer Milo, Elad Berkman, and Raviv Pryluk. 2024. "Robust CATE Estimation Using Novel Ensemble Methods." *arXiv preprint arXiv:2407.03690*.

Mackey, Lester, Vasilis Syrgkanis, and Ilias Zadik. 2018. "Orthogonal machine learning: Power and limitations." In *International Conference on Machine Learning*, pp. 3375-3383. PMLR.

Nie, Xinkun, and Stefan Wager. 2021. "Quasi-oracle estimation of heterogeneous treatment effects." *Biometrika,* 108(2): 299-319.

Oprescu, Miruna, Vasilis Syrgkanis, and Zhiwei Steven Wu. 2019. "Orthogonal random forest for causal inference." In *International Conference on Machine Learning*, pp. 4932-4941. PMLR.

Palensky, Peter, and Dietmar Dietrich. 2011. "Demand side management: Demand response, intelligent energy systems, and smart loads." *IEEE transactions on National Academyrmatics* 7(3): 381-388.

Park, Junhyung, Uri Shalit, Bernhard Schölkopf, and Krikamol Muandet. 2021. "Conditional distributional treatment effect with kernel conditional mean embeddings and u-statistic regression." In *International conference on machine learning*, pp. 8401-8412. PMLR.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70 (1): 41–55.

Rosenbaum, Paul, and Donald Rubin. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *The American Statistician,* 39(1): 33-38.

Sailor, David. 2001. "Relating residential and commercial sector electricity loads to climate—evaluating state level sensitivities and vulnerabilities." *Energy,* 26(7): 645-657.

Souto, Hugo Gobato, and Francisco Louzada Neto. 2024. "K-Fold Causal BART for CATE Estimation." *arXiv preprint arXiv:2409.05665*.

Strittmatter, Anthony. 2023. "What is the value added by using causal machine learning methods in a welfare experiment evaluation?" *Labour Economics,* 84: 102412.

Stuart, Elizabeth. 2010. "Matching methods for causal inference: A review and a look forward." *Statistical Science: A Review Journal of the Institute of Mathematical Statistics,* 25(1): 1.

Wager, Stefan, and Susan Athey. 2018. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association,* 113(523): 1228-1242.

Wu, Guojun, Ge Song, Xiaoxiang Lv, Shikai Luo, Chengchun Shi, and Hongtu Zhu. 2023. "DNet: distributional network for distributional individualized treatment effects." In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 5215-5224.

Zadik, Ilias, Lester Mackey, and Vasilis Syrgkanis. 2018. "Orthogonal machine learning: Power and limitations." In *International Conference on Machine Learning*, pp. 5723-5731.

Zhou, Tianhui, William Carson, and David Carlson. 2022. "Estimating potential outcome distributions with collaborating causal networks." *Transactions on Machine Learning Research*.

Appendix

Table A1: Covariate balance around zone-specific CDD threshold using SMD

	75 Bandwidth		100 Bandwidth		125 Bandwidth	
Variables	T-Stat	SMD	T-Stat	SMD	T-Stat	SMD
Marine Climate Zone	0.441	0.018	-0.239	-0.009	-0.502	-0.016
Hot and Mixed Dry Climate Zone	-0.635	-0.026	-1.644*	-0.059	-0.014	0.000
Mixed Humid Climate Zone	-1.660*	-0.068	-1.272	-0.046	-2.493*	-0.081
Hot Humid Climate Zone	-0.824	-0.033	-0.456	-0.016	0.125	0.004
Central AC Type	1.383	0.056	1.639*	0.059	2.769*	0.089
Evaporative Coolers AC Type	0.487	0.020	0.298	0.011	-0.101	-0.003
20k-39k income range	-0.003	0.000	0.229	0.008	-0.375	-0.012
40k-59k income range	-0.727	-0.029	-0.265	-0.009	0.058	0.002
60k-79k income range	2.715*	0.109	2.954*	0.105	2.747*	0.088
80k-99k income range	0.373	0.015	0.029	0.001	0.235	0.008
100k and above income range	-1.963*	-0.079	-2.121*	-0.076	-2.277*	-0.073
Asian Households	1.020	0.041	1.980*	0.071	1.666*	0.053
Multiracial Households	-0.313	-0.013	-0.387	-0.014	0.172	0.006
Native American Households	-0.309	-0.013	-1.228	-0.044	-0.466	-0.015
Pacific Islander Households	-1.415	-0.060	-1.415	-0.053	-1.733*	-0.057
White Households	-2.014*	-0.081	-1.717*	-0.061	-2.120*	-0.068
Mobile Homes	-0.330	-0.013	-1.200	-0.043	-1.032	-0.033
Single Family Homes	-1.978*	-0.080	-2.141*	-0.077	-0.688	-0.022
Presence of Insulation	-1.027	-0.041	-0.573	-0.021	-0.962	-0.031
Log Cooled SqFt	-0.555	-0.022	-0.476	-0.017	0.590	0.019
Number of rooms	-0.867	-0.035	-1.081	-0.039	0.149	0.005
Household Size	-1.965*	-0.079	-2.241*	-0.080	-1.277	-0.041
Bulbs on for 1-4hrs	0.482	0.019	0.148	0.005	0.497	0.016
Log HDD	-0.631	-0.026	-1.503	-0.054	-4.037*	-0.130

Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A1: Robustness analysis – treatment effects based on 30-year average CDD

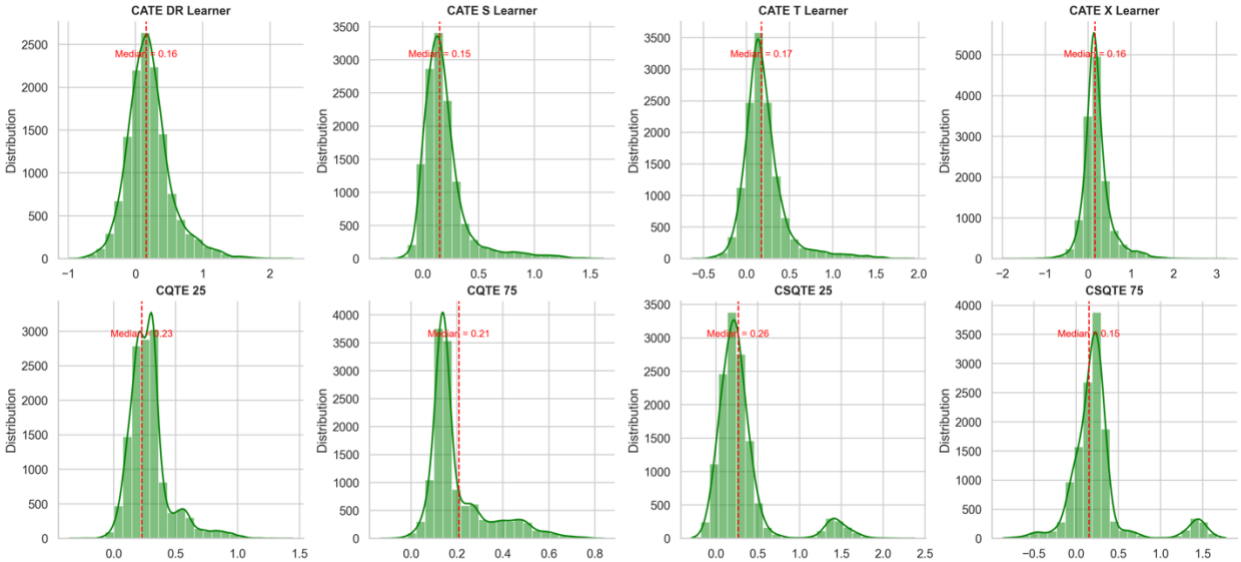


Table A2: Heterogeneous effects under 30-Year average CDD treatment threshold

	CATE (S Learner)	CATE (T Learner)	CATE (X Learner)	CATE (DR Learner)	CQTE (25)	CQTE (75)	CSQTE (25)	CSQTE (75)
Panel A: Air Conditioner Type								
Central AC Type	0.040*** (0.004)	0.053*** (0.008)	0.036*** (0.006)	0.066*** (0.007)	0.068*** (0.008)	-0.015** (0.007)	0.123*** (0.003)	-0.041*** (0.003)
Evaporative coolers	0.072*** (0.012)	0.044** (0.022)	0.048*** (0.016)	0.093*** (0.026)	0.063*** (0.026)	-0.458*** (0.024)	0.096*** (0.008)	0.070*** (0.009)
Constant	0.160*** (0.004)	0.163*** (0.007)	0.178*** (0.006)	0.152*** (0.007)	0.258*** (0.006)	0.277*** (0.006)	0.187*** (0.003)	0.231*** (0.003)
R^2	0.006	0.004	0.003	0.007	0.047	0.039	0.082	0.030
N	13198	13198	13198	13198	13198	13198	13198	13198
Panel B: Income Group								
$20k–$39k	0.026*** (0.007)	0.039*** (0.012)	0.034*** (0.009)	0.030*** (0.011)	0.091*** (0.013)	0.053*** (0.012)	0.038*** (0.006)	0.018*** (0.004)
$40k–$59k	0.047*** (0.007)	0.052*** (0.012)	0.049*** (0.009)	0.049*** (0.011)	0.179*** (0.013)	0.120*** (0.012)	0.042*** (0.006)	0.044*** (0.005)
$60k–$79k	0.030*** (0.008)	0.030** (0.013)	0.032*** (0.010)	0.051*** (0.012)	0.291*** (0.014)	0.084*** (0.013)	0.056*** (0.006)	0.023*** (0.005)
$80k–$99k	0.034*** (0.007)	0.032*** (0.012)	0.041*** (0.009)	0.028** (0.011)	0.072*** (0.013)	0.087*** (0.012)	0.037*** (0.006)	0.013*** (0.005)
$100k+	0.043*** (0.006)	0.053*** (0.011)	0.053*** (0.008)	0.042*** (0.009)	0.057*** (0.011)	0.031*** (0.011)	0.041*** (0.005)	0.017*** (0.004)
Constant	0.160*** (0.005)	0.166*** (0.009)	0.169*** (0.007)	0.170*** (0.008)	0.223*** (0.010)	0.195*** (0.009)	0.248*** (0.004)	0.182*** (0.004)
R^2	0.004	0.002	0.004	0.002	0.044	0.011	0.007	0.008
N	13198	13198	13198	13198	13198	13198	13198	13198
Panel C: Household Race								
Asian	0.093*** (0.016)	0.148*** (0.026)	0.144*** (0.019)	0.111*** (0.023)	0.318*** (0.028)	0.219*** (0.026)	0.080*** (0.013)	0.080*** (0.010)
Multiracial	0.083*** (0.018)	0.122*** (0.026)	0.125*** (0.021)	0.096*** (0.026)	0.036 (0.034)	0.126*** (0.032)	0.089*** (0.013)	0.066*** (0.010)
Native American	0.087*** (0.027)	0.108** (0.044)	0.139*** (0.039)	-0.040 (0.050)	-0.165*** (0.054)	0.233*** (0.050)	0.046** (0.021)	0.099*** (0.018)
Pacific Islander	0.125 (0.087)	0.137 (0.111)	0.126 (0.089)	-0.072 (0.134)	0.086*** (0.144)	0.136** (0.133)	0.153** (0.062)	0.082* (0.043)
White	-0.002 (0.007)	-0.005 (0.011)	0.007 (0.008)	-0.014 (0.011)	-0.018 (0.012)	0.061*** (0.011)	0.008 (0.006)	0.035*** (0.004)
Constant	0.189*** (0.007)	0.201*** (0.011)	0.193*** (0.008)	0.212*** (0.011)	0.330*** (0.011)	0.189*** (0.010)	0.274*** (0.006)	0.166*** (0.004)
R^2	0.011	0.010	0.015	0.007	0.032	0.012	0.012	0.013
N	13198	13198	13198	13198	13198	13198	13198	13198
Panel D: Housing Type								
Mobile Home	0.055*** (0.008)	0.147*** (0.017)	0.095*** (0.012)	0.072*** (0.015)	-0.034** (0.017)	0.020 (0.016)	0.060*** (0.007)	0.045*** (0.007)
Single family Home	0.036*** (0.005)	0.051*** (0.009)	0.025*** (0.007)	0.021*** (0.008)	0.106*** (0.009)	0.147*** (0.009)	0.054*** (0.004)	0.030*** (0.003)
Constant	0.162*** (0.005)	0.158*** (0.008)	0.184*** (0.006)	0.185*** (0.008)	0.246*** (0.009)	0.138*** (0.008)	0.241*** (0.004)	0.176*** (0.003)
R^2	0.005	0.008	0.005	0.002	0.015	0.027	0.015	0.008
N	13198	13198	13198	13198	13198	13198	13198	13198

Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix A1: Algorithmic Details for CATE Estimation

This appendix provides the full algorithmic exposition for the meta-learners used to estimate CATE in Chapter 3.

A1.1 S-Learner

The S-learner estimates a single response function:

$$\tau(X_i, A_i) = \mathrm{E}[Y_i | X_i, A_i] \qquad (A3.1)$$

The CATE is then obtained as:

$$\widehat{\tau_s}(X_i) = \hat{\tau}(X_i, 1) - \hat{\tau}(X_i, 0) \qquad (A3.2)$$

While computationally efficient, this approach constrains treated and control outcomes to share a common functional form.

A1.2 T-Learner

The T-learner estimates separate outcome models for treated and control households:

$$\tau_0(X_i) = \mathrm{E}[Y_i(0) | X_i], \quad \tau_1(X_i) = E[Y_i(1) | X_i] \qquad (A3.3)$$

The CATE estimate is:

$$\widehat{\tau_T}(X_i) = \widehat{\tau_1}(X_i) - \widehat{\tau_0}(X_i) \qquad (A3.4)$$

This approach allows distinct functional forms across treatment states but may suffer from higher variance when treatment groups are imbalanced.

A1.3 X-Learner

The X-learner improves efficiency under treatment imbalance by imputing individual-level treatment effects. First, counterfactual outcomes are imputed:

$$\widetilde{D_1} = Y_i(1) - \widehat{\tau_0}(X_i), \qquad \widetilde{D_0} = \widehat{\tau_1}(X_i) - Y_i(0) \qquad (A3.5)$$

Second-stage models estimate conditional expectations of these imputed effects. The final CATE is a weighted average:

$$\widehat{\tau_x}(X_i) = g(X_i)\widehat{\tau_0}(X_i) + \big(1 - g(X_i)\big)\widehat{\tau_1}(X_i) \qquad (A3.6)$$

where $g(X_i)$ reflects relative group sizes or propensity scores.

A1.4 Doubly Robust (DR) Learner

The DR-learner combines outcome regression and propensity score estimation. Let

$$e(X_i) = Pr(A_i = 1|X_i)$$

denote the propensity score. The pseudo-outcome is constructed as:

$$\widetilde{D_\iota} = A_i\left(\frac{Y_i - \widehat{\tau_0}(X_i)}{e(X_i)}\right) + (1 - A_i)\left(\frac{\widehat{\tau_1}(X_i) - Y_i}{1 - e(X_i)}\right) \qquad (A3.7)$$

Regressing $\widetilde{D_\iota}$ on X_i yields the DR CATE:

$$\widehat{\tau_{DR}}(X_i) = E\big[\widetilde{D_\iota} \,\big|\, X_i\big]$$

This estimator is consistent if either the outcome model or the propensity score model is correctly specified, making it particularly attractive in observational settings.

All learners are implemented with cross-fitting to maintain orthogonality between nuisance estimation and treatment effect estimation. Gradient boosting and random forest models are used as base learners. Standard errors are obtained via repeated sample splitting and aggregation.

Appendix A2: Algorithmic Details for CDTE Estimation

This appendix provides the algorithmic exposition underlying the estimation of CDTEs summarized in Section 3.4.3.

A2.1 Pseudo-Outcome Construction

Following Kallus and Oprescu (2023), CDTE estimation relies on a pseudo-outcome that corrects for bias in naïve plug-in estimators. Let $Z = (Y, A, X)$ denote the observed data, and let $e(X) = P(A = 1|X)$ denote the propensity score. Define nuisance functions $\aleph_A(X)$ and $h_A(X)$ estimated separately for treated $(A = 1)$ and control $(A = 0)$ groups. The CDTE pseudo-outcome is given by:

$$\psi(Z; e, \aleph, h) = \aleph_1(X) - \aleph_0(X) - \frac{A - e(X)}{e(X)\big(1 - e(X)\big)}\varrho\big(Y, \aleph_A(X), h_A(X)\big) \qquad (A3.8)$$

Under standard regularity and overlap conditions, the pseudo-outcome satisfies:

$$E[\psi(Z; e, \aleph, h)|X] = CDTE(X) \qquad (A3.9)$$

A2.2 Conditional Quantile Treatment Effects

For CQTE estimation, the conditional quantile function $\aleph_A(X) = q_A(X; \tau)$ is obtained by solving:

$$E\big[\varrho\big(Y, \aleph_A(X)\big)|X, A = a\big] = 0$$

where $\varrho(\cdot)$ denotes the quantile moment condition.

The density of the outcome at the conditional quantile enters through:

$$h_A(X) = \frac{1}{f_{Y|X,A=a}\big(q_a(X; \tau)\big)} \qquad (A3.10)$$

with $f_{Y|X,A=a}(\cdot)$ estimated using flexible conditional density methods.

A2.3 Conditional Super-Quantile Treatment Effects

For CSQTEs, estimation proceeds in two steps. First, the conditional quantile $q_A(X; \tau)$ is estimated for each treatment group. Second, the super-quantile is computed as:

$$\mu_a(X;\tau) = E[(1-\tau)^{-1} Y I[Y \geq q_a(X;\tau)] \mid X, A = a] \qquad (A3.11)$$

This approach captures tail-average effects beyond the specified quantile and is particularly informative for assessing extreme electricity consumption responses.

A2.4 Cross-Fitting and Final Estimation

All nuisance components including propensity scores, conditional quantiles, densities, and super-quantiles are estimated using cross-fitting to ensure independence between nuisance estimation and pseudo-outcome construction. The final CQTE, and CSQTE estimates are obtained by regressing the pseudo-outcomes on covariates X, yielding consistent and asymptotically normal estimates of distributional treatment effects.

Appendix A3: Algorithmic Implementation Details

A3.1 Cross-Fitting and Pseudo-Outcome Construction

All treatment effect estimators employ KKK-fold cross-fitting to decouple nuisance parameter estimation from treatment effect estimation. For each fold, nuisance components—including the propensity score and conditional outcome functions—are estimated on training folds and used to construct pseudo-outcomes on held-out folds. This procedure ensures Neyman orthogonality and guards against bias arising from regularization or overfitting.

A3.2 Conditional Distributional Treatment Effects

CQTE and CSQTE estimation follows the pseudo-outcome regression framework of Kallus and Oprescu (2023). Conditional quantiles are first estimated separately by treatment status, after which pseudo-outcomes are constructed to target quantile- and tail-specific treatment effects. CSQTEs extend this approach by estimating expected outcomes beyond the quantile threshold, capturing tail risk behavior.

A3.3 CATE Meta-Learners

CATEs are estimated using S-, T-, X-, and DR-learners. Outcome models are implemented using Gradient Boosting Regressors, while propensity scores are estimated using Random Forest Classifiers. The DR-learner combines outcome regression and propensity weighting to achieve double robustness under correct specification of either component.

A3.4 Software Implementation

All estimation procedures are implemented in Python. CATE learners follow the framework of Künzel et al. (2019), while distributional estimators are based on the CDTE learner of Kallus and Oprescu (2023). Hyperparameters are selected via cross-validation, and HC3-robust standard errors are reported for all projection regressions.