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.

1.0 Introduction

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

The objective of this study is to analyze how climate-induced cooling demand differentially impacts household electricity consumption across population subgroups and levels of energy use. Using a causal machine learning framework grounded in the potential outcomes approach, the study aims to uncover heterogeneous responses to cooling demand, quantify how these effects vary across the consumption distribution, and identify which household groups are most affected by rising cooling needs. By focusing on both the average and distributional effects, the analysis provides evidence relevant for designing equitable and adaptive energy policies.

Several studies including Auffhammer & Aroonruengsawat (2011) and Deschênes & Greenstone (2011) document the strong links between temperature and residential electricity consumption, including responses during extreme heat. Yet mean-based approaches can mask important dispersion in responses across households with different technologies, building stocks, and resources. Moreover, program evaluations and pricing pilots show wide heterogeneity in demand responses, underscoring the need to understand who bears the largest marginal burden when cooling demand rises (Reiss & White 2005; Ito 2014; Knittel & Stolper 2022). These findings motivate the use of flexible causal inference methods capable of capturing complex and heterogeneous behavioral responses to climatic variation.

This study adopts a causal machine learning (ML) framework to move beyond descriptive associations. Unlike unconditional quantile regression, which relies on functional-form assumptions and influence function projections (Firpo et al. 2009), causal ML is grounded in the potential outcome framework, enabling estimation of treatment effects that vary across households

and the consumption distribution. Although increasingly applied in other policy contexts (Athey et al. 2019; Knittel & Stolper 2022), its use in climate–energy analysis remains limited. Methodologically, causal ML framework flexibly capture nonlinear, high-dimensional interactions (Wager & Athey 2018; Athey et al. 2019), while orthogonalization and cross-fitting minimize bias and overfitting, ensuring valid inference under standard causal assumptions (Chernozhukov et al. 2018; Nie & Wager 2021; Künzel et al. 2019). These properties are particularly valuable with survey-based energy data, where the effects of climate exposure are mediated by diverse demographic and technological factors.

The study uses 2020 RECS microdata for the analysis. Based on geographic variation in cooling exposure, the treatment is defined as a binary indicator equal to one if a household's CDD exceeds the mean CDD of its climate zone, enabling within-zone comparisons. This zone-specific normalization controls for structural climatic differences and enables meaningful within-zone comparisons of cooling-related electricity consumption. Balance checks on pre-treatment covariates are conducted within ±75, ±100, and ±125 CDD bandwidths using standardized mean differences. Robustness is assessed via two alternative treatment definitions: first, households above the 75th percentile of CDD within their climate zone, and second, households whose 30-year average CDD exceeds the zone-specific mean. Estimated CATE, CQTE, and CSQTE effects are projected via OLS onto household characteristics, linking nonparametric results to observable traits and showing subgroup patterns for targeted energy policies.

The findings show that higher cooling degree days significantly increase household electricity consumption by roughly 32–34% on average with disproportionately larger effects among lower-consuming, minority, and middle-income households. These results highlight widening energy-use inequalities under climate stress and underscore the importance of targeted efficiency and adaptation policies that address household-level vulnerabilities.

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 utilizes the 2020 wave of the Residential Energy Consumption Survey (RECS), which provides nationally representative data on household energy use. As in Gbadegesin (2025), continuous variables with right-skewed distributions such as electricity consumption and cooling degree days (CDD) are log-transformed to allow for elasticity-based interpretations. Data cleaning, trimming, and weight construction follow the procedures previously outlined in Gbadegesin (2025). The 2020 RECS include 18,496 households, offering a substantially larger sample than the 5,686 households in 2015. After applying data cleaning and trimming criteria, the final analytical sample consists of 13,198 observations.

Focusing on a single wave rather than pooling multiple years is supported in the literature, as repeated cross-sectional designs like RECS help avoid issues related to attrition, respondent conditioning, and inconsistent follow-up that often complicate panel or pooled analyses (Lipps, 2021). Moreover, combining survey years may obscure time-specific differences in sampling (Lebo & Weber, 2015). In addition, the exclusive use of the 2020 wave not only takes advantage of the larger sample size but also ensures analytical consistency within a more recent and homogeneous context. This focus allows for clearer identification of heterogeneous effects and serves as a validation of the distributional patterns identified in the pooled UQR results from Gbadegesin (2025). Additionally, the 2020 data reflect the most recent climate conditions and household energy behaviors, particularly in light of increasing extreme heat events and ongoing shifts in residential cooling technologies.

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

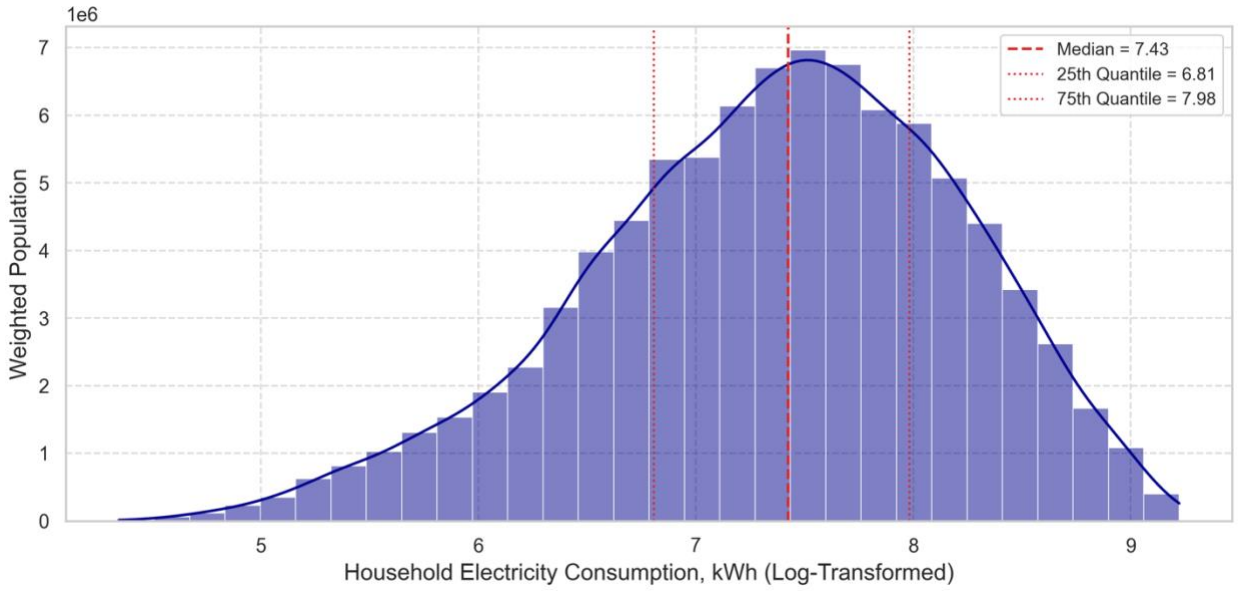
3.3 Outcome Variable of Interest

The primary outcome variable in this study is annual household electricity consumption for space cooling, measured in kilowatt-hours (kWh). This variable captures the quantity of electricity households use to maintain thermal comfort during warmer periods and serves as a direct measure of behavioral response to cooling energy demand. According to U.S. EIA, it is an estimate of electricity used for space cooling via central AC, window/wall units, or evaporative coolers—derived from survey data, load shapes, climate, and appliance characteristics; excludes non-system fans/blowers. Unlike Gbadegesin (2025), which also considered electricity expenditure, the current analysis isolates consumption in physical units to better assess heterogeneity in usage patterns without conflating behavioral effects with price variation or billing structures. The measure is derived from calibrated RECS estimates and reflects household-level cooling loads across different system types and climate conditions.

In the energy economics literature, 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 (Andor et al. 2021; Wang et al. 2023; and U.S. EIA 2022). In contrast, electricity expenditures are influenced by income, price responsiveness, and affordability constraints—factors that may reflect financial conditions rather than true consumption behavior (Ali et al. 2021). As a result, reliance on expenditure measures may obscure the analysis of usage heterogeneity and bias elasticity estimates.

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 Treatment Variable of Interest

To motivate the construction of the treatment variable, this study first examines the distribution and descriptive statistics of CDD using both 2020 observations and 30-year averages. Figure 3.2 presents the distribution of CDD across U.S. households where 2020 CDD values are unimodal and slightly right-skewed, with a median of 7.20, equivalent to approximately 1,341 CDDs, and an interquartile range from 6.86 (954 CDDs) to 7.63 (2,062 CDDs). Panel B show a similar distribution but a slightly narrower interquartile range from 6.66 (778 CDDs) to 7.55 (1,889 CDDs). Table 3 complements these visualizations by reporting disaggregated summary statistics by climate zone. U.S. households in the Hot Humid zone experienced the highest average CDD in 2020 (3,310), followed by those in Hot & Mixed Dry (2,254) and Mixed Humid (1,501) zones, whereas households in the Cold & Very Cold and Marine zones experienced much lower averages of 896 and 637, respectively. The same relative pattern is observed in the 30-year average data, though with slightly lower values. These descriptive findings underscore the substantial geographic heterogeneity in cooling demand and reinforce the importance of applying climate zone-specific thresholds or percentile-based classifications when constructing treatment variables to capture high cooling demand.

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

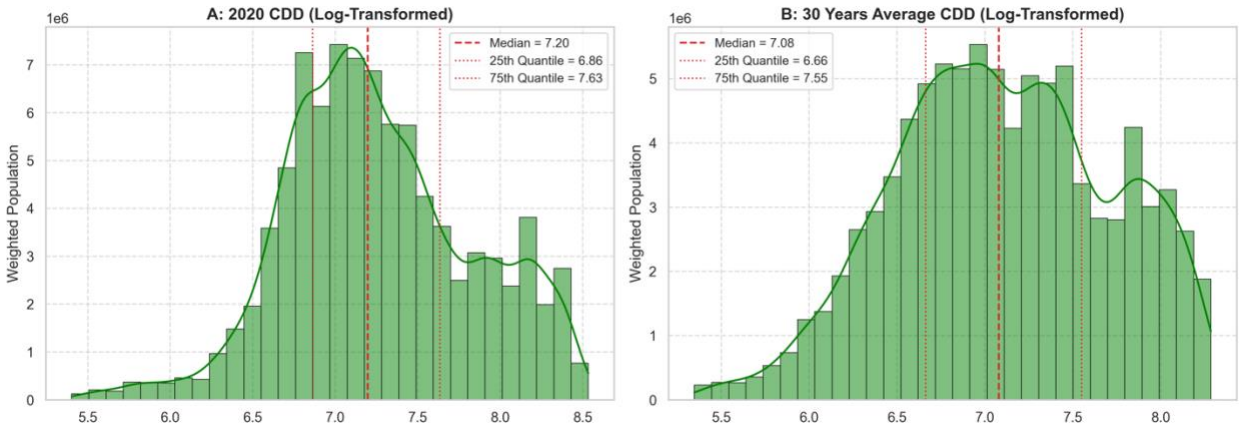


Table 3: Weighted descriptive statistics of cooling degree days by climate zones

	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

Based on the observed geographic heterogeneity across climate zones, the treatment variable is defined as a binary indicator of high cooling demand. Specifically, 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. Formally, the treatment indicator is defined as:

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

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.

The choice of SMD as the primary balance metric follows best-practice guidance in the causal-inference literature, which recommends standardized differences over significance testing because they are sample-size invariant and directly quantify imbalance (Austin 2009; Stuart 2010; Rosenbaum and Rubin 1985). This approach aligns with the broader tradition of balance diagnostics in matching and regression-discontinuity designs in observational studies (Ho et al. 2007; Imbens and Rubin 2015; Jacobsen 2019; Cattaneo and Titiunik 2022). By demonstrating that SMDs remain below the conventional 0.10 threshold across bandwidths, the study provides credible evidence that exceeding the climate-zone-specific CDD threshold generates plausibly exogenous variation in cooling demand.

The results, summarized in Table 4, indicate strong covariate balance across all three bandwidths, with nearly all SMDs falling below the conventional 0.1 threshold. The narrower ±75 CDD bandwidth achieves the strongest balance, with virtually all covariates well below 0.1. The ±100 CDD bandwidth maintains similarly strong balance, though a few variables such as mid-level income ($60,000–$79,000) and White race indicators show statistically significant differences at the 10% level, while still remaining under the 0.1 threshold. The wider ±125 CDD bandwidth introduces slightly more variation, with most covariates balanced but HDD reaching an SMD of 0.13. These results demonstrate that covariate balance is preserved across bandwidths, underscoring the robustness of the design to different window choices.

Table 4: Covariate balance around zone-specific CDD threshold using SMD

Covariates	75 Bandwidth		100 Bandwidth		125 Bandwidth	
	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.

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 ($CDD_{z,2020}^{(75)}$). This alternative specification identifies households experiencing particularly intense cooling needs relative to local climatic norms. Formally, this is defined as:

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

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 ($\bar{C}DD_{z,30yr}$).

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

The treatment indicator A_{iz} thus captures whether household i in climate zone z experiences high cooling demand, normalized to local climatic conditions. This variable form the basis for the causal analysis that follows. In subsequent sections, the empirical models quantify how exposure to high cooling demand causally affects household electricity consumption. Section 3.7 develops this framework formally by estimating both conditional mean and distributional treatment effects using causal machine learning methods.

3.5 Control & Moderating Variables

To ensure robust identification of the effect of cooling demand on household electricity consumption, the estimation strategy incorporates the full set of control variables introduced in Gbadegesin (2025). These include continuous and categorical variables accounting for climatic exposure (e.g., climate zone, HDD), housing characteristics (e.g., cooled square footage, number of rooms), household composition (e.g., household size), and socioeconomic and behavioral factors (e.g., income brackets, AC equipment type, lighting usage). All control variables are operationalized as described in Gbadegesin (2025).

To explore potential sources of heterogeneity, the CATE, CQTE, and CSQTE estimates are projected on a set of covariates including income group, AC equipment type, household race, and housing unit type. This allows the analysis to assess how the effect of cooling demand varies across different household and structural characteristics. 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 that builds upon and extends the UQR framework. While UQR provided important insights into how CDD affects different points of the electricity consumption distribution, it is inherently limited in causal interpretation due to its reliance on the Recentered Influence Function and its focus on unconditional quantiles. Moreover, UQR does not fully address treatment effect heterogeneity or characterize responses across the entire outcome distribution. To overcome these limitations, this study adopts the CATE and CDTE which leverages the potential outcomes model to estimate electricity consumption under treated and untreated states, $Y_i(1)$ and $Y_i(0)$, thereby enabling the estimation of causal effects conditional on observed covariates.

Having defined the treatment variable (A_i) and relevant control and moderating variables (X_i), the empirical strategy proceeds in multiple stages. Section 3.6.1 introduces the estimation of CATE using meta-learners (S-, T-, X-, and DR-learners), which allow flexible, nonparametric modeling of heterogeneous treatment effects via algorithms such as random forests and kernel regressions. Section 3.6.2 extends this framework to CDTE (CQTE and CSQTE), which quantifies differences across the conditional distributions of electricity consumption under treatment and control, enabling the study of not only mean responses but also tail risks and extreme outcomes. The inclusion of CSQTE is particularly valuable for identifying high-risk, energy-burdened households.

To further explain the sources of heterogeneity, each estimated treatment effect—CATE, CQTE, and CSQTE—is subsequently projected onto interpretable household and building characteristics using OLS. This projection step ensures asymptotically valid inference and facilitates the identification of characteristics most strongly associated with cooling demand. Compared to the parametric, interaction-based approach in Gbadegesin (2025), this framework provides a more flexible and robust lens into how and for whom cooling demand driven by climate conditions influences household electricity consumption.

3.6.1 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 use under these two scenarios, conditional on household characteristics (X_i). Formally, CATE is expressed as:

$$CATE(X_i) = E[Y_i(1) - Y_i(0)|X_i] \qquad (3.1)$$

where, X_i includes household-level covariates such as climate zone, 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:

$$Y_i = A_i Y_i(1) + (1 - A_i) Y_i(0) \qquad (3.2)$$

Under the assumptions of unconfoundedness ($Y(a) \perp A|X$) and overlap ($0 < P(A = 1|X) < 1$), treatment effects can be estimated using observed data.

To estimate CATE, this study implements four meta-learners: the S-learner, T-learner, X-learner, and DR-learner, following the methodology outlined in Künzel et al. (2019). The S-learner treats the treatment indicator, A_i, as a standard feature in the regression model, estimating a single response function:

$$\tau(X_i, A_i) = \mathrm{E}[Y_i|X_i, A_i] \qquad (3.3)$$

using a supervised learning model trained on the entire dataset. The estimated function, $\hat{\tau}(X_i, A_i)$, is then used to compute the CATE as:

$$\widehat{\tau_s}(X_i) = \hat{\tau}(X_i, 1) - \hat{\tau}(X_i, 0) \qquad (3.4)$$

This approach is computationally efficient, as it fits only one model, but it assumes the same functional form for both treated and control groups, which may be restrictive in some cases.

The T-learner takes a two-step approach, estimating separate models for the treated and control groups. First, the control response function is estimated as:

$$\tau_0(X_i) = \mathrm{E}[Y_i(0)|X_i] \qquad (3.5)$$

using a supervised learning model trained only on the control group ($A_i = 0$). Similarly, the treatment response function is estimated as:

$$\tau_1(X_i) = \mathrm{E}[Y_i(1)|X_i] \qquad (3.6)$$

using another supervised learning model trained only on the treated group ($A_i = 1$). The CATE estimate from the T-learner is then:

$$\widehat{\tau_T}(X_i) = \widehat{\tau_1}(X_i) - \widehat{\tau_0}(X_i) \qquad (3.7)$$

This method allows for different functional forms in the treated and control groups, making it more flexible than the S-learner. However, when the sample sizes for one group are significantly smaller, model estimates can suffer from high variance.

The X-learner extends the T-learner by leveraging imputed treatment effects to improve estimation, especially when the treatment and control groups are highly unbalanced. The X-learner follows three main steps, first estimate the response functions using equation 3.3. For the treated group ($A_i = 1$) and control group ($A_i = 0$), impute the counterfactual control and treated outcome, respectively:

$$\widetilde{D_1} = Y_i(1) - \widehat{\tau_0}(X_i) \qquad (3.8)$$

$$\widetilde{D_0} = \widehat{\tau_1}(X_i) - Y_i(0) \qquad (3.9)$$

CATE is estimated using a weighted average:

$$\widehat{\tau_x}(X_i) = g(X_i)\widehat{\tau_0}(X_i) + \big(1 - g(X_i)\big)\widehat{\tau_1}(X_i) \qquad (3.10)$$

where $g(X_i)$ is the weighting function, typically chosen as:

$$g(X_i) = \frac{n_0}{n_0 + n_1}$$

The DR Learner combines propensity score modeling and outcome regression to reduce bias in CATE estimation. It is particularly useful in observational settings where treatment assignment may depend on observed covariates. It first estimates the propensity score:

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

using a classification model. Then, separate outcome models for treated and control groups are fitted:

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

Using these estimates, the pseudo-outcome is obtained as:

$$\widetilde{D_i} = A_i\left(\frac{Y_i - \widehat{\tau_0}(X_i)}{e(X_i)}\right) + (1 - A_i)\left(\frac{\widehat{\tau_1}(X_i) - Y_i}{1 - e(X_i)}\right) \qquad (3.11)$$

The pseudo-outcome $\widetilde{D_i}$ is then used as the response variable in a regression model to estimate CATE.

$$\widehat{\tau_{DR}}(X_i) = E\left[\widetilde{D_i} \mid X_i\right]$$

3.6.2 Conditional distributional treatment effects specification

Extending the same household-level treatment indicator A_i defined in Section 3.4, the analysis adopts a distributional perspective to capture how high cooling demand affects the entire conditional distribution of electricity consumption. Building on Kallus and Oprescu (2023), the CDTE framework extends beyond mean treatment effects to quantify differences in specific statistics (e.g., quantiles, super-quantiles) of the conditional distributions of electricity use under treatment $\left(F_{Y(1)|X}\right)$ and control $\left(F_{Y(0)|X}\right)$ conditions. The CDTE is formally expressed as:

$$CDTE(X) = \aleph^*\left(F_{Y(1)|X}\right) - \aleph^*\left(F_{Y(0)|X}\right) \qquad (3.12)$$

where $\aleph^*(F)$ represents a statistic derived from a moment condition:

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

This 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:

$$CQTE(X;\tau) = q\left(F_{Y(1)|X};\tau\right) - q\left(F_{Y(0)|X;\tau}\right) \qquad (3.13)$$

where $q(F;\tau)$ represents the quantile at level τ. This measure provides insights into how the treatment affects households at different points in the outcome distribution, such as those at the median or in the upper or lower tails.

The second measure is the Conditional Super-Quantile Treatment Effect (CSQTE), which extends the analysis by capturing the average outcomes beyond a specified quantile, focusing on tail risks. Unlike CQTEs, which focus on specific thresholds, CSQTEs provide a broader view of the risk profile by considering the average effects in the extreme tail of the distribution. It is expressed as:

$$CSQTE(X;\tau) = \mu\big(F_{Y(1)|X};\tau\big) - \mu\big(F_{Y(0)|X;\tau}\big) \qquad (3.14)$$

where $\mu(F;\tau)$ represents the expected value of outcomes exceeding the τ-quantile.

To estimate CDTEs, the study adopts the pseudo-outcome regression approach introduced by Kallus and Oprescu (2023). This method corrects for potential bias in naive plug-in estimators and ensures robust CDTE estimation. The pseudo-outcome for CDTE learning is defined as:

$$\psi(Z;e,\aleph,h) = \aleph_1(X) - \aleph_0(X) - \frac{A - e(X)}{e(X)\big(1-e(X)\big)}\varrho\big(Y,\aleph_A(X),h_A(X)\big) \qquad (3.15)$$

where $e(X) = P(A=1|X)$, is the propensity score; $\aleph_A(X)$ and $h_A(X)$ are nuisance functions estimated separately for treated $(A=1)$ and control $(A=0)$ groups; $\varrho(Y,\aleph,h)$ is the moment condition linking observed outcomes and the estimated statistics.

The pseudo-outcome satisfies equation 17, ensuring consistency when regressed on X.

$$E[\psi(Z;e,\aleph,h)|X] = CDTE(X) \qquad (3.16)$$

For CQTEs, the conditional quantile function is estimated by solving:

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

For CSQTEs, the super-quantile is estimated in two steps. First, the conditional quantile $q_a(X;\tau)$ is estimated for each treatment group. Then, the expected value of outcomes exceeding the quantile is calculated. This process can be implemented by splitting the training data, estimating the quantile on one subset, and computing tail averages on the other.

$$\mu_a(X;\tau) = E[(1-\tau)^{-1}YI[Y \geq q_a(X;\tau)] \mid X, A=a] \qquad (3.17)$$

For CQTEs, the reciprocal of the density at the conditional quantile is required:

$$h_A(X) = \frac{1}{f_{Y|X,A=a}\big(q_a(X;\tau)\big)} \tag{3.18}$$

where $f_{Y|X,A=a}$ is the density of Y given X and $A = a$.

3.6.3 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 capture the heterogeneous treatment effects associated with each covariate. The projection is performed separately for each group to identify the characteristics most strongly associated with variation in treatment effects across the different 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:

$$\widehat{\tau_\iota} = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \varepsilon_i \tag{3.19}$$

where $\widehat{\tau_\iota}$ 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:

$$\beta^* = \underset{\beta \epsilon Rp}{arg\ min} \,||\, CDTE - \phi(X)^T \beta \,|| \tag{3.20}$$

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:

$$\sqrt{n}\left(\hat{\beta} - \beta^*\right) \xrightarrow{d} N(0, \Sigma) \qquad (3.21)$$

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.

3.6.4 Estimation approach

The estimation of CDTEs is carried out using the CDTE learner developed by Kallus and Oprescu (2023), which employs cross-fitting to ensure independence between nuisance function estimation and pseudo-outcome construction. First, the data is divided into K folds. For each fold k, nuisance functions including the propensity score, $e(X)$, conditional quantiles $\aleph_A(X)$, and density estimates $h_A(X)$ are estimated using data from all folds except the k-th fold. This approach ensures that the estimation of nuisance functions is performed independently of the samples used for final pseudo-outcome calculations, maintaining an orthogonal structure to the treatment effect estimation. These estimates are then used to construct a pseudo-outcome for each observation. After computing the pseudo-outcomes across all folds, the final step involves estimating the conditional expectation of the pseudo-outcomes given the covariates. This step aggregates the information in the pseudo-outcomes to estimate the CDTE, which captures how the treatment effect varies across different levels of the covariates.

For the estimation of CQTE, the process involves constructing models for quantile treatment effects at specific quantiles, utilizing a combination of Random Forest Quantile Regressors, conditional density models, and a linear regression learner for final estimation. The Random Forest Quantile Regressor is employed to estimate conditional quantiles of the outcome distribution given covariates (X) and treatment (A), capturing the treatment effect at the 25th and 75th percentiles of the outcome distribution (Y). Additionally, a conditional density model, represented by a Random Forest Regressor, is used to approximate the conditional density function of the covariates, and a kernel-based function is included to model fine-grained dependencies. These components feed into the final CQTE estimation model, represented by a linear regression with robust covariance

estimation. The pseudo-outcomes are computed using these models and then aggregated to estimate the CQTE.

For the estimation of CSQTE, two distinct models were utilized, the Random Forest Quantile Regressor and the Kernel Superquantile Regressor. The Random Forest Quantile Regressor estimates conditional quantiles of the outcome distribution given covariates (X), capturing treatment effects at specific percentiles of the outcome distribution. In contrast, the Kernel Superquantile Regressor extends this analysis to estimate conditional super quantiles, reflecting expected outcome values in the tail beyond the specified quantile threshold. This distinction is critical for assessing tail behavior, such as high risks or rewards, and is key to understanding distributional effects.

For CATE estimation, this study employs multiple meta-learners, including the S-learner, T-learner, X-learner, and DR Learner, leveraging Gradient Boosting Regressors and Random Forest Classifiers to estimate heterogeneous treatment effects. The T-learner estimates separate outcome models for the treated and control groups, capturing differential responses by applying Gradient Boosting Regression to both groups before computing the treatment effect as the difference in predicted outcomes. The S-learner incorporates the treatment indicator as a feature in a single Gradient Boosting Regressor, estimating potential outcomes in a unified model. The X-learner extends the T-learner by imputing treatment effects for each group and refining estimates with a second-stage model, incorporating a Random Forest Classifier for propensity score estimation. Finally, the DR Learner combines outcome modeling with propensity score weighting, where a Gradient Boosting Regressor is used for both the outcome and pseudo-treatment models, while a Random Forest Classifier estimates the propensity score, ensuring robustness against misspecification in either component. These approaches effectively capture non-linear relationships and high-dimensional interactions among covariates, enhancing the flexibility and accuracy of CATE estimation.

The final stage of the estimation process involved projecting the estimated effects onto selected covariates using OLS. This step provides interpretable coefficients for CATE, CQTE, and CSQTE by regressing the estimated effects onto a subset of key categorical variables. Specifically, the projection included variables such as household air conditioning type (e.g., central air

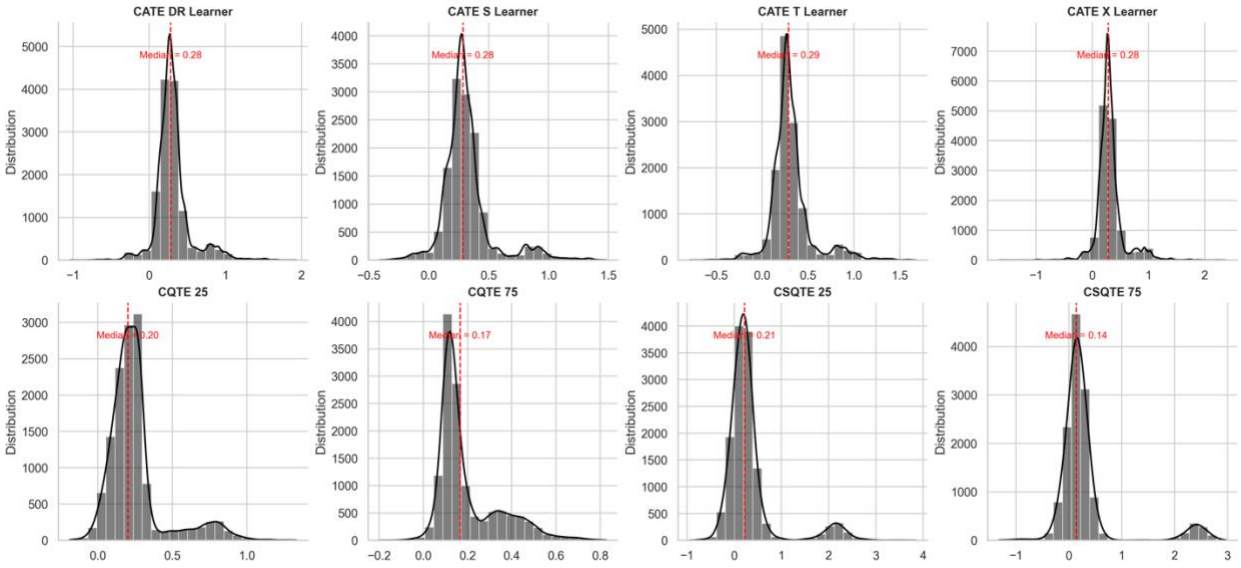
conditioning, unitary air conditioning, evaporative coolers), household race (e.g., Asian, Black, Multiracial, Native American, Pacific Islander, White), income group (e.g., below $20,000; $20,000–$39,000; $40,000–$59,000; $60,000–$79,000; $80,000–$99,000; and $100,000 or more), and household building type (e.g., mobile homes, single-family homes, apartment homes). All estimation procedures in this study are implemented in Python, with CATE learners following the modules of Künzel et al. (2019) and CDTE estimations (CQTE and CSQTE) based on the modules developed by Kallus and Oprescu (2023).

4.0 Empirical Findings

4.1 Electricity consumption response to cooling demand

Figure 3 presents the treatment effect estimates for each estimator across the full sample, including four CATE learners (T, S, X, and DR Learners), and the CQTE and CSQTE estimates at the 25th and 75th quantiles. For the CATE estimates, the median values are approximately 0.28–0.29 across all learners. This implies that households experiencing higher-than-average cooling demand consume approximately 32.2% to 33.6% more electricity compared to their counterparts within the same climate zone. The CQTE estimates show heterogeneous effect across the outcome distribution. At the 25th quantile, the median CQTE is 0.20, suggesting a 22.1% increase in consumption among lower-consuming households. At the 75th quantile, the effect is slightly lower, with a median of 0.17, implying an 18.5% increase. The pattern remains consistent for the CSQTE (tail expectation) estimates but shows a slightly flatter gradient. The median CSQTE is 0.21 at the 25th quantile and 0.14 at the 75th quantile, implying electricity consumption increases of approximately 23.4% and 15.0%**,** respectively. These distributional estimates underscore how cooling demand disproportionately burdens more vulnerable households.

Figure 3: Distribution of estimated effects under increased cooling demand



4.2 Heterogeneous effects by socioeconomic and housing characteristics

Table 6 presents results from an OLS projection of CATE, CQTE and CSQTE on air conditioning system types, with unitary AC types serving as the reference group. Based on the CATE models, total effects for central AC users is about 0.32 which translates to approximately 37.7% higher electricity consumption relative to unitary AC households of about 32%. At the 25th quantile, the CQTE estimate for central AC users is 0.337, corresponding to a 40.1% increase, while at the 75th quantile the effect drops to 34.3%. The CSQTE estimates show a similar pattern: households with central AC experience a 31.6% increase at the 25th quantile and 19.6% at the 75th quantile. In comparison, households using evaporative coolers face even greater increases under CATE models, with total effects ranging from approximately 47.6% to 53.6% higher electricity consumption compared to unitary AC users. While the CQTE effects for evaporative coolers are lower (25.6% and 30.6 at 25th and 75th quantile, respectively), the CSQTE estimates are substantial: households see a 25.9% increase at the 25th percentile and a 37.7% increase at the 75th percentile.

The result shows a nuanced pattern emerges across models: in the CATE projections, households using evaporative coolers consistently show the highest average increase in electricity consumption, followed by those with central AC and unitary AC systems. In contrast, the CQTE projection indicate that central AC households exhibit stronger responsiveness to cooling demand than evaporative cooler households at both lower and upper quantiles. However, this pattern

reverses in the CSQTE, where households with evaporative coolers face a steeper increase in electricity use at higher levels of consumption compared to those with central AC systems.

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.								

Relative to households earning below $20,000 which experience approximately a 35% increase in electricity consumption under elevated cooling demand, the OLS projections show (Table 7) a non-monotonic pattern of heterogeneity across income groups. For the CATE outcome, the total effect for households earning $20,000–$39,000 is approximately 41%, 39% for those earning $40,000–$59,000, and a peak of 43% for the $60,000–$79,000 group. In contrast, the effect declines to around 27% for the $80,000–$99,000 group, and returns to 37% for those earning $100,000 and above. This suggests that while all income groups experience substantial increases in consumption under high cooling demand, the mid-income households, particularly those between $60,000 and $79,000, exhibit the highest responsiveness.

Additional result from the CQTE outcome reinforce this pattern. At the 25th percentile of the consumption distribution, households earning $40,000–$59,000 experience the largest increase, approximately 67%, compared to 38% for the lowest income group. Meanwhile, at the 75th percentile, the $60,000–$99,000 income brackets show the highest increases, each around 55%, while the lowest and highest income groups experience lower impacts—32% for $20,000–$39,000 and 29% for $100,000 and above. CSQTE outcome, which capture tail-averaged effects, are relatively flatter across groups. In the lower tail, effects range between 23%–31%, while in the upper tail, they cluster between 20%–24%. These results indicate that while the burden of increased cooling demand is broadly shared, middle-income households bear disproportionately higher effects in terms of average electricity consumption. However, at the lower and upper ends

of the consumption distribution, the effects appear more diffuse, with different income segments exhibiting higher responsiveness depending on the quantile or tail considered.

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 on household race (Table 8), using Black households as the reference group, a clear pattern emerges in how cooling demand impacts electricity consumption. In CATE projection, Asian households have total effects that are nearly identical to Black households—roughly 39–41% increases, with differences statistically insignificant. In contrast, White households consistently show lower responsiveness, around 36%. For the CQTE regression, at the 25th percentile, Asian households show a 65% increase in electricity consumption under high cooling demand, nearly double the 34% increase for White households. At the 75th percentile, Asians still exhibit higher responsiveness (38%), though both Asians and Whites remain below the Black baseline (33%). The tail-expectation regression models reinforce these disparities: for the lower tail, Asians have a 38% response versus Whites at 28%, and in the upper tail, Asians remain higher at 25%, relative to 22% for Whites.

Compared to Black households, Multiracial households do not show statistically significant differences in CATE projection but exhibit a significant increase in household electricity consumption due to cooling demand in the CSQTE projection—approximately 41.6% at the 25th percentile. Native American households experience significantly higher increases in electricity use

under CATE projection (up to 57.8% based on the S Learner), and despite showing notably lower effects in the CQTE projections, they display substantial increases in consumption in the CSQTE estimates—about 43.9% and 30.7% at the 25th and 75th percentiles, respectively. Pacific Islander households exhibit the largest increases in electricity consumption due to cooling demand across most models, with estimates as high as 70.1% (X Learner) and CSQTE effects of 58.7% at the 25th percentile, indicating a strong sensitivity to cooling needs relative to Black households. These findings consistently indicate that while Black households face a substantial burden from increased cooling needs, other minority groups—especially Native American and Pacific Islander households—tend to bear even greater increases in electricity consumption, particularly in the tails of the distribution.

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 show that, compared to apartment homes, households in Mobile Homes experience approximately 43.6% to 46.1% higher electricity consumption due to increased cooling demand, based on the CATE projections. Households in Single Family Homes also exhibit elevated consumption, though to a slightly lesser extent (38.5%) but still higher than apartment households. In the CQTE regression, Single Family Homes show around 44.7% higher consumption at the 25th percentile and 42.9% at the 75th percentile, indicating strong responsiveness to cooling demand across the distribution. By contrast, Mobile Homes display a slight increase of roughly 19.9% at

the 25th percentile and 17.4% at the 75th percentile, though these effects are not statistically significant. For the CSQTE projections, Single Family Homes show about a 31.3% increase in electricity use at the lower tail and 22.1% at the upper tail, while Mobile Homes reflect approximately 22% at the lower tail and no significant difference at the upper tail.

These results indicate that while the burden of increased cooling demand is broadly shared across housing types, households in single-family homes bear disproportionately higher effects in terms of average electricity consumption compared to apartment and mobile homes. However, at the lower and upper quantile of the consumption distribution, the effects appear more diffuse, with single-family households showing stronger responsiveness overall, while mobile-home households exhibit marginal or statistically insignificant effects depending on the quantile or tail considered.

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.

4.3 Robustness analysis using alternative cooling degree day thresholds

Figures 4 and 5 provide robustness checks of Figure 3.3 findings. Figure 3.4 redefines treatment as exposure to extreme heat, where 2020 CDD exceeds the 75th percentile within each climate zone. Under this specification, CATE estimates across learners range from 0.13 to 0.20, implying consumption increases of 13.9% to 22.1% for those facing unusually high cooling needs. CQTE medians remain elevated, at 0.19 (≈ 20.9%) for the 25th quantile and 0.16 (≈ 17.4%) for the 75th, with CSQTE estimates reinforcing this pattern at 22.1% and 17.4%, respectively. These results confirm that even households with moderate or low baseline electricity use are significantly impacted by intense heat exposure.

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

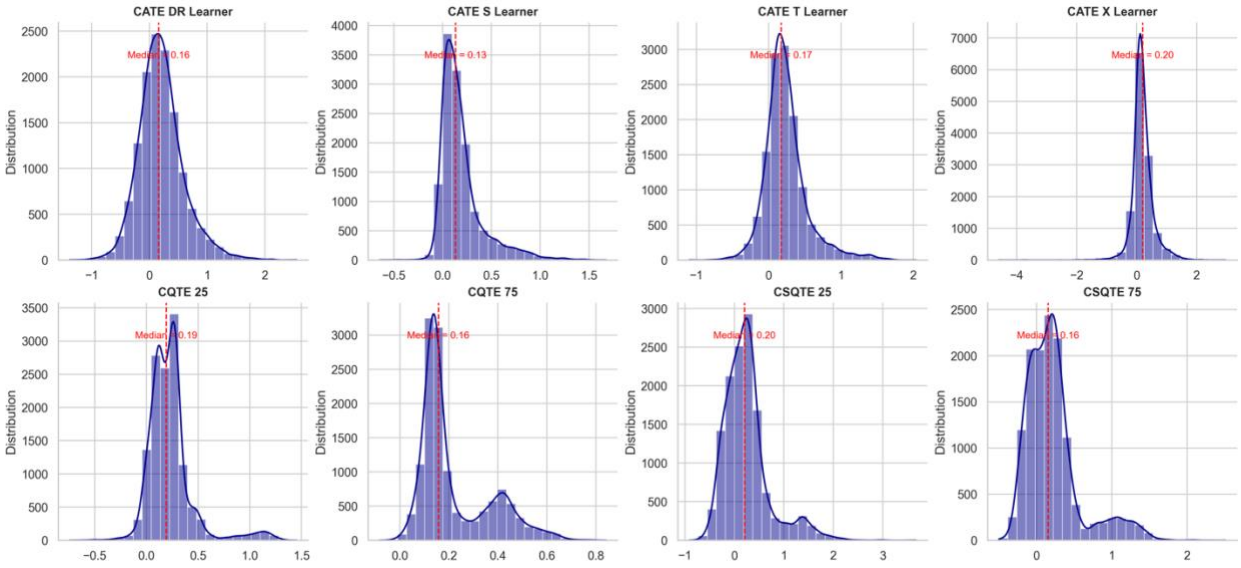
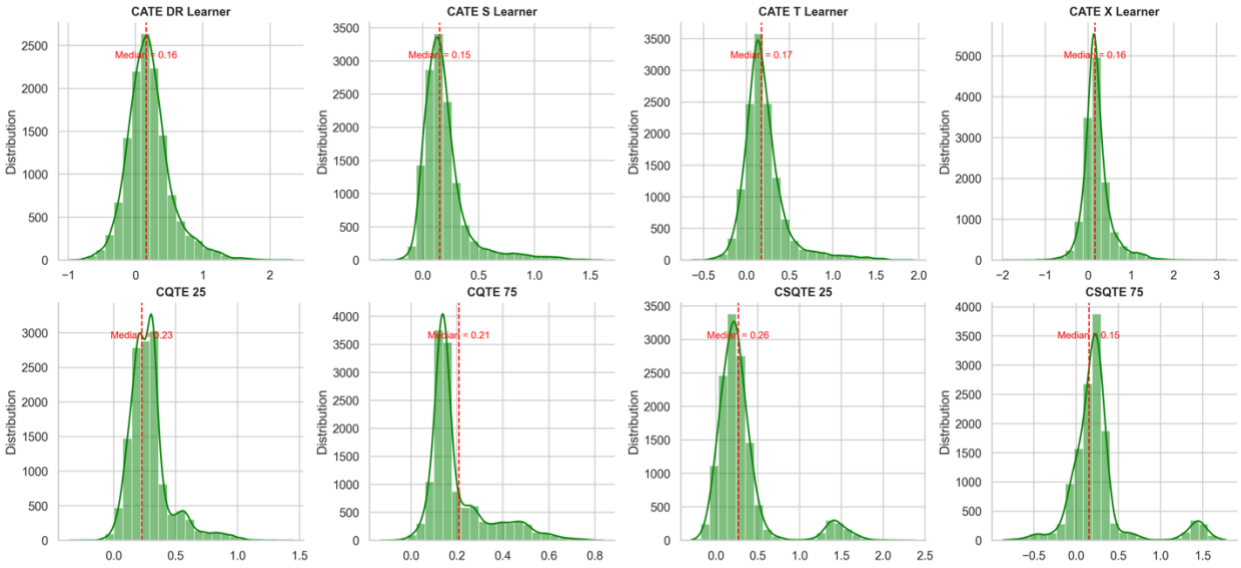


Figure 5 tests the long-term climatic robustness of the results by redefining treatment using 30-year average CDD. Households are considered treated if their long-run average CDD exceeds the mean of their climate zone. CATE medians remain positive and stable (0.15–0.17), implying 16.2% to 18.5% increases in electricity consumption among long-term heat-exposed households. CQTE results show even stronger effects among lower-consuming households, with a 25th quantile median of 0.23 (≈ 25.9% increase) compared to 0.21 (≈ 23.4%) at the 75th. CSQTE estimates also reinforce this distributional pattern, with the highest median effect of 0.26 (≈ 29.8%) occurring in the lower tail. Overall, this robustness check supports the main findings and demonstrates their stability under alternative treatment definitions.

Figure 5: Robustness analysis – treatment effects based on 30-year average CDD



Under the alternative treatment definition based on the 75th percentile of CDD within each climate zone, the OLS projections of CATE, CQTE, and CSQTE (Table 10) reinforce the main findings, preserving the core patterns while amplifying certain distributional effects. In Panel A, central AC households continue to exhibit higher electricity consumption than the unitary AC baseline, with a notably stronger lower-tail response (approximately 28% at the 25th quantile), whereas evaporative cooler households remain distinctly upper-tail intensive, with increases of about 40% at the 75th quantile. In Panel B, the hump-shaped income gradient is maintained, with households earning $60k–$79k again displaying the largest responses, including CQTE effects at the 25th quantile exceeding 60%, while higher-income groups show reduced upper-tail effects.

Panel C confirms pronounced racial heterogeneity: Black, Asian, Native American, and Pacific Islander households register particularly large tail responses, with upper-quantile effects often exceeding 45%, whereas White households exhibit more moderate and consistent increases. This reinforces the disproportionate electricity consumption burden on minority groups. In Panel D, housing type asymmetries persist, with mobile homes showing upper-tail increase (above 25% at the 75th quantile) and single-family homes concentrating higher consumption among lower-use households (31% at the 25th quantile) while recording a lower effect in the upper quantiles.

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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

When the treatment is redefined using 30-year average CDD, the OLS projection results (Table 11) reinforce the baseline findings. By air conditioning type, central AC households continue to consume more than unitary AC users, with a pronounced lower-quantile response (34% at the

25th-quantile CQTE) even as upper-tail CSQTE effects decline (21%). Evaporative coolers retain large distributional sensitivity, with a sizable lower-tail CQTE effect (36.8%). The income gradient remains hump-shaped, with middle-income households, particularly those earning $60k–$79k, again exhibiting the strongest distributional response (68% at the 25th-quantile CQTE), while upper-income groups show more moderate increases, reinforcing that the highest marginal burden falls on the middle of the income distribution.

Racial heterogeneity also persists and intensifies as minority groups, especially Blacks, Asian, Native American and Pacific Islander households bear substantial burden from increased cooling demand. For example, Asian households show the largest lower-tail effect (76% at the 25th-quantile CQTE), while Native American and Multiracial households record strong upper-tail responses (60–65% at the 75th-quantile CQTE). Housing results also aligns with mobile homes exhibiting asymmetric responses where lower-tail consumption is above baseline (27.9% at the 25th-quantile CQTE compared to 23.6% apartment homes) while single-family homes record 38.9% increase at the 25th quantile and 31.6% at the 75th quantile and 31.1% at the 25th quantile and 21.0% at the 75th quantile for CSQTE model, confirming that single-family homes bear disproportionately higher effects in terms of household electricity consumption.

Table 11: 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***	0.053***	0.036***	0.066***	0.068***	-0.015**	0.123***	-0.041***
	(0.004)	(0.008)	(0.006)	(0.007)	(0.008)	(0.007)	(0.003)	(0.003)
Evaporative coolers	0.072***	0.044**	0.048***	0.093***	0.063***	-0.458***	0.096***	0.070***
	(0.012)	(0.022)	(0.016)	(0.026)	(0.026)	(0.024)	(0.008)	(0.009)
Constant	0.160***	0.163***	0.178***	0.152***	0.258***	0.277***	0.187***	0.231***
	(0.004)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.003)	(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***	0.039***	0.034***	0.030***	0.091***	0.053***	0.038***	0.018***
	(0.007)	(0.012)	(0.009)	(0.011)	(0.013)	(0.012)	(0.006)	(0.004)
$40k–$59k	0.047***	0.052***	0.049***	0.049***	0.179***	0.120***	0.042***	0.044***
	(0.007)	(0.012)	(0.009)	(0.011)	(0.013)	(0.012)	(0.006)	(0.005)
$60k–$79k	0.030***	0.030**	0.032***	0.051***	0.291***	0.084***	0.056***	0.023***
	(0.008)	(0.013)	(0.010)	(0.012)	(0.014)	(0.013)	(0.006)	(0.005)
$80k–$99k	0.034***	0.032***	0.041***	0.028**	0.072***	0.087***	0.037***	0.013***
	(0.007)	(0.012)	(0.009)	(0.011)	(0.013)	(0.012)	(0.006)	(0.005)
$100k+	0.043***	0.053***	0.053***	0.042***	0.057***	0.031***	0.041***	0.017***
	(0.006)	(0.011)	(0.008)	(0.009)	(0.011)	(0.011)	(0.005)	(0.004)
Constant	0.160***	0.166***	0.169***	0.170***	0.223***	0.195***	0.248***	0.182***
	(0.005)	(0.009)	(0.007)	(0.008)	(0.010)	(0.009)	(0.004)	(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***	0.148***	0.144***	0.111***	0.318***	0.219***	0.080***	0.080***
	(0.016)	(0.026)	(0.019)	(0.023)	(0.028)	(0.026)	(0.013)	(0.010)
Multiracial	0.083***	0.122***	0.125***	0.096***	0.036	0.126***	0.089***	0.066***
	(0.018)	(0.026)	(0.021)	(0.026)	(0.034)	(0.032)	(0.013)	(0.010)
Native American	0.087***	0.108**	0.139***	-0.040	-0.165***	0.233***	0.046**	0.099***
	(0.027)	(0.044)	(0.039)	(0.050)	(0.054)	(0.050)	(0.021)	(0.018)
Pacific Islander	0.125	0.137	0.126	-0.072	0.086***	0.136**	0.153**	0.082*
	(0.087)	(0.111)	(0.089)	(0.134)	(0.144)	(0.133)	(0.062)	(0.043)
White	-0.002	-0.005	0.007	-0.014	-0.018	0.061***	0.008	0.035***
	(0.007)	(0.011)	(0.008)	(0.011)	(0.012)	(0.011)	(0.006)	(0.004)
Constant	0.189***	0.201***	0.193***	0.212***	0.330***	0.189***	0.274***	0.166***
	(0.007)	(0.011)	(0.008)	(0.011)	(0.011)	(0.010)	(0.006)	(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***	0.147***	0.095***	0.072***	-0.034**	0.020	0.060***	0.045***
	(0.008)	(0.017)	(0.012)	(0.015)	(0.017)	(0.016)	(0.007)	(0.007)
Single family Home	0.036***	0.051***	0.025***	0.021***	0.106***	0.147***	0.054***	0.030***
	(0.005)	(0.009)	(0.007)	(0.008)	(0.009)	(0.009)	(0.004)	(0.003)
Constant	0.162***	0.158***	0.184***	0.185***	0.246***	0.138***	0.241***	0.176***
	(0.005)	(0.008)	(0.006)	(0.008)	(0.009)	(0.008)	(0.004)	(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: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.4 Comparison with UQR Results

The CQTE OLS projection results align closely with the UQR interaction estimates by producing the similar ranking of responsiveness across key subgroups while providing quantile-specific causal effects. For cooling technologies, both approaches show central AC households as more responsive than unitary AC users, but the CQTE results refine this pattern by showing that the excess burden is concentrated in the lower tail (40% at the 25th percentile) and diminishes toward the upper tail, whereas evaporative coolers display smaller, more balanced effects across tails (26%–31%). For income, CQTE confirms the heavier upper-tail burdens for mid-to-upper income households (55% at the 75th percentile for $60k–$99k) observed in the UQR models but also show that the largest lower-tail effect occurs among $40k–$59k households (67%), indicating a non-monotonic pattern. Racial disparities are similarly reinforced as minority groups maintain elevated responses across tails, and White households display more moderate increases. Housing type patterns remain consistent, with single-family homes exhibiting strong, significant effects at both tails (45% and 43%, respectively), while mobile homes show smaller and statistically insignificant changes.

The CSQTE projections reinforce the heterogeneity observed in the CQTE and UQR models while identifying where the most severe shocks occur under elevated cooling demand. In contrast to CQTE, CSQTE indicates that evaporative coolers bear the steepest upper-tail increases (37.7% at the 75th percentile), while central AC systems maintain higher lower-tail burdens (31.6% at the 25th percentile). Across income groups, tail effects are more evenly distributed (20%–31% across both tails), suggesting that extreme consumption risk is shared across the income spectrum even where average or median effects differ. Racial disparities remain pronounced and for housing, single-family homes display the largest and most consistent tail burdens (31% at the lower tail), while mobile homes, despite larger average effects in CATE, show weaker and less consistent extreme responses.

The CATE projections provide the average treatment effects across the full distribution, serving as a baseline for interpreting the quantile-specific CQTE and tail-averaged CSQTE results. Both central AC and evaporative cooler households consume substantially more electricity than unitary AC users on average (≈37.7% and up to ≈53.6%, respectively), underscoring their higher baseline heat-driven electricity burdens. Income effects are non-monotonic, peaking among $60k–$79k

households (≈43%) and remaining substantial for both the lowest- and highest-income groups. Racial and ethnic differences are marked: Native American households show significantly higher average effects (up to ≈57.8%), and Pacific Islander households face the largest average burdens (up to ≈70.1%), while Asian and White households record lower averages relative to Black households. Housing patterns are consistent with distributional findings, as mobile homes (≈44–46%) and single-family homes (≈38.5%) exhibit much higher average increases than apartments, reflecting the greater structural exposure of detached housing to heat-related electricity demand.

5.0 Conclusion

This study examined the heterogeneous causal effects of climate-driven cooling demand on household electricity consumption using 2020 Residential Energy Consumption Survey (RECS) data in a causal machine learning framework grounded in the potential outcomes approach. The analysis proceeds in three parts. First, CATE are estimated to identify how the causal impact of high versus low cooling demand varies across households with different characteristics. Second, CQTE and CSQTE are used to assess how treatment effects differ across the electricity consumption distribution, capturing heterogeneity among low- and high-consuming households. Third, the estimated treatment effects are projected onto key household attributes including income level, race or ethnicity, air conditioning type, and housing structure to determine which groups are most affected by rising cooling needs.

The findings highlight the substantial increase in household electricity consumption associated with elevated cooling demand. Across the full sample, the CATE estimates suggest that households experiencing higher-than-average cooling demand consume roughly 32% to 34% more electricity than their peers within the same climate zone. The CQTE and CSQTE estimates further show that the burden is unevenly distributed: lower-consuming households experience more pronounced increases—about 21% to 23%—than those at the 75th percentile, where effects range from 15% to 19%.

Robustness checks confirm the stability of the results under alternative definitions of extreme heat exposure. When defining treatment based on exposure to CDD levels above the 75th percentile, CATE effects range between 13.9% and 22.1%. CQTE and CSQTE estimates remain elevated, with consumption increases of approximately 21% at the lower end of the distribution and 17.4%

at the upper end. A second robustness specification, using 30-year average CDD as treatment, yields similarly consistent results. Under this long-run climate exposure definition, electricity use increases by 16–18.5% on average (CATE), and up to 26–30% in the lower tail (CQTE and CSQTE).

Further disaggregation based on OLS projection by air conditioning system types show important differences in electricity consumption responses. Households with evaporative coolers consistently exhibit the highest average increases, with CATE projection suggesting up to 53.6% higher consumption relative to unitary AC users. Central AC users also show strong effects, averaging 37.7%. However, the CQTE results indicate greater responsiveness for central AC users at both the 25th and 75th percentiles compared to evaporative coolers. This pattern reverses under the CSQTE models, where evaporative coolers are associated with significantly larger increases, particularly at the upper tail (up to 37.7%).

Income-based heterogeneity uncovers a non-linear pattern in electricity consumption due to cooling demand. While all groups experience notable increases, middle-income households ($60,000–$79,000) show the highest responsiveness in average consumption, with effects peaking at 43%. The CQTE model highlights the largest effects in the 25th percentile for middle-income households (up to 67%) and in the 75th percentile for those in the mid-to-upper brackets (around 55%). The CSQTE estimates are more evenly distributed but still suggest slightly higher impacts among middle-income groups. These results indicate that while the burden of cooling demand is widespread, mid-income households disproportionately bear the highest average impact, though the tails of the distribution reflect more diffuse effects.

Finally, projections by household race and housing type confirm additional heterogeneity. Relative to Black households, Asian households exhibit similar average consumption increases under CATE but significantly higher effects at the lower and upper quantiles (CQTE and CSQTE). In contrast, White households show lower responsiveness across most models. Native American and Pacific Islander households face the steepest increases, particularly in the distribution tails, with some effects exceeding 70%. Regarding housing, single-family homes consistently show higher responsiveness to cooling demand compared to apartments and mobile homes. While mobile

homes exhibit large average increases under CATE (up to 46%), these effects are not statistically significant in the CQTE or CSQTE projections.

The two robustness checks—redefining treatment as 2020 CDD above the climate-zone 75th percentile and using 30-year average CDD above the zone mean—confirm the baseline OLS projection results across the CATE, CQTE, and CSQTE specifications. In CATE, central AC households show persistently higher consumption, while CQTE show strong lower-tail responses (28%–34% at the 25th quantile) and CSQTE indicate flatter or declining upper tails. Evaporative coolers display more variable distributional effects, with CQTE capturing large impacts in either the upper or lower tail depending on the treatment definition, and CSQTE showing concentrated upper-tail burdens. The income gradient remains robustly hump-shaped, with the $60k–$79k group showing the largest effects—CQTE at the 25th quantile exceeds 60%—while higher-income households consistently exhibit smaller upper-tail responses in CQTE and CSQTE. Racial disparities persist across both specifications, with minority groups such as Black, Asian, Native American, and Pacific Islander households bearing substantial burden of electricity consumption following increased cooling demand. Housing type patterns are also robust, with single-family households disproportionately more affected relative to apartments and mobile homes.

The CQTE projection results align closely with the UQR interaction estimates, producing the similar responsiveness while adding causal quantile-specific effect. Both approaches show central AC households as more responsive than unitary AC users, but CQTE refines this by showing that the excess burden is concentrated in the lower tail, whereas evaporative coolers have smaller, more balanced effects across tails. The income gradient observed in UQR is also confirmed, with CQTE highlighting heavier upper-tail burdens for mid-to-upper income households and a pronounced lower-tail peak for $40k–$59k households, indicating a non-monotonic pattern. These patterns are reinforced by the CSQTE projections, which show that extreme consumption burdens follow similar subgroup hierarchies but are more evenly distributed across income groups, with certain technologies (e.g., evaporative coolers) and racial groups (e.g., Pacific Islanders, Native Americans) exhibiting the steepest tail responses. The CATE results further support these findings by confirming that the subgroups identified as most responsive in the quantile models—such as households with central AC or evaporative coolers, middle-income earners, and certain minority

racial groups—also face the largest average heat-related electricity burdens across the full distribution.

The findings from this study underscore that climate-driven cooling demand imposes significant and uneven burdens on household electricity use, with impacts varying by consumption level, technology type, income, race, and housing structure. This heterogeneity underscores that one-size-fits-all energy efficiency or climate adaptation policies may fail to protect the most vulnerable subgroups. For example, middle-income households, minority racial groups, and households with evaporative coolers or central AC face disproportionately high average or tail burdens, suggesting the need for targeted policies that are sensitive to subgroup-specific vulnerabilities. The disparities suggest that rising cooling needs could exacerbate existing inequalities in energy costs, particularly under extreme heat conditions.

Methodologically, this study advances the growing literature on applying causal machine learning techniques in energy and climate research. By integrating CATE, CQTE, and CSQTE estimators with OLS projection, and employing robust treatment assignment strategies, the study generates effects that are plausibly causal. This framework enhances the internal validity of the results and provides a replicable template for future empirical research on climate–energy interactions using high-dimensional observational data. The consistency of results across alternative treatment definitions further reinforces their credibility for informing both short-term heat adaptation measures and long-term climate resilience strategies. In practical terms, policies such as energy efficiency retrofits, targeted subsidies for high-efficiency cooling systems, and tailored demand-side management programs can play a critical role in alleviating the disproportionate burden borne by high-risk households.

References

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.

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.

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.

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, 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.

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*.