Climate Adaptive Response Estimation: Short And Long Run Impacts Of Climate Change On Residential Electricity and Natural Gas Consumption Using Big Data*

Maximilian Auffhammer
University of California, Berkeley & NBER

December 16, 2016

Abstract

The widespread rollout of air conditioning has resulted in massive positive impacts on human health in both developing and developed countries over the past century. A warming world will further increase demand for cooling both through increased operation of existing air conditioners and the adoption of new ones, while at the same time decreasing the demand for heating. This paper provides an empirical approach to separately quantifying the intensive and extensive margin response for cooling without having to make explicit assumptions about the current day or future penetration of air conditioners. We compare the estimated increases in electricity demand from cooling to projected decreases in natural gas demand from lowered heating demand. Exploiting a panel dataset containing close to two billion electricity and natural gas bills for California, we estimate the electricity and gas temperature response of residential consumption at the zip code level. Using the most recent set of IPCC climate models, we simulate future consumption by letting households vary consumption along their current response function as well as letting the electricity response function change using an econometrically estimated second stage. We show that reductions in natural gas demand more than offset any climate driven increases in electricity consumption. We further show that failing to account for extensive margin adjustment in electricity demand leads to a significant underestimate of future impacts.

Keywords: Climate Change, Electricity Demand, Natural Gas Demand, Heating, Cooling
JEL Codes: Q41, Q54

*207 Giannini Hall #3310, Berkeley, CA. Tel: (510) 643-5472. Email: auffhammer@berkeley.edu. I thank the California Energy Commission's PIER program for generous funding. Seminar participants at Stanford University, UC Irvine, Iowa State University, UCSB, UIUC, Lawrence Berkeley Labs, Arizona State University and Resources for the Future have provided valuable feedback. This draft is preliminary and may not be cited yet.

1. Introduction

Anthropogenic emissions of greenhouse gases are projected to significantly alter the global climate over the current century and beyond. The Intergovernmental Panel on Climate Change (IPCC) projects that the average global surface temperature will rise by between 1 and 3.7 ° Celsius by the end of the century. If significant mitigation efforts are not undertaken in the near future, warming is likely to be at the higher end of that range. This shift in the mean of the global surface temperature distribution will be accompanied by significant increases in the frequency and intensity of extreme heat events (IPCC AR5, 2013). One of the three Integrated Assessment Models used in the calculation of the Social Cost of Carbon concludes that increased space cooling is the major driver in Global Climate Damages (Rose et al, 2016).

Air conditioning is the main adaptation mechanism to hotter climate open to humans and has been shown to be an incredibly effective strategy to mitigate the negative health impacts of hot days. Barreca et al. (2016) show that in the United States the mortality effect of a very hot day decreased by roughly 80% between 1900-1959 and 1960-2004. It is important to consider that the historical trajectory of air conditioner installation has been mainly driven by growth in incomes and the falling prices of the units and electricity over the past century. As a consequence, the cross sectional distribution of air conditioners shows higher penetration levels in hotter and wealthier areas (EIA, 2011). There is a dearth of causal estimates of empirically calibrated damage functions, which quantify the relationship of higher temperatures and electricity consumption from space cooling.

One dimension of this adaptation response to the higher incidence of extreme heat days arising from global climate change, will be the more frequent operation of existing air conditioning equipment, which we will refer to as the intensive margin adjustment for the remainder of the paper. The other response will be the installation of air conditioners in areas that currently see little penetration of this equipment. We will refer to this dimension of adaptation as the extensive margin adjustment. While there are a number of papers attempting to quantify the intensive margin adjustment, a full characterization of the extensive margin at fine enough levels of aggregation to be

useful to planners is extremely difficult, as data on installed air conditioners over time and space are not available for most areas. One exception is Davis and Gertler (2015) who observe the universe of electricity bills for Mexico and survey of air conditioner ownership for 27,000 households. Their simulations display the massive importance of accounting for the extensive margin adjustments.

The main innovation of this paper is that we lay out a simple method to estimate both the intensive and extensive margin impacts of climate change on electricity consumption when one *does not have data installed capital*. In a first stage, we estimate the temperature response function of household electricity consumption at a fine level of spatial aggregation - the five-digit ZIP code level. These response functions allow us to examine how the intensive margin adjustment ("increased usage of existing equipment") varies across 1,235 ZIP codes in California in our sample. Estimation at this fine level is made possible by the fact that we observe almost 1 billion electricity bills which represent 80 percent of California's household over a decade.

In a second stage, we explain cross-sectional variation in these "first stage" estimated slopes of each ZIP code's temperature response function as a function of long run weather ("climate") and income. Technically speaking, in this "second stage" we regress the slope of each ZIP code's temperature response function in different temperature bins on income and summer climate. We therefore separate the impact of income on temperature response from the direct effect of climate. The estimated marginal effect of climate on the slope of the response function allows us to capture extensive margin adjustments to long run changes in climate. We use downscaled predictions from 18 of the IPCC's most recent climate models to simulate future household electricity consumption at the ZIP code level under climate change, taking into account both intensive ("first stage") and extensive margin ("second stage") adjustments. We then compare the projected increases in electricity consumption to climate driven reductions in natural gas consumption due to warmer winters, which are derived from a "first stage" intensive margin only regression (as one would not expect people to install more or less heating equipment due to milder winters). We show that in the case of California's residential sector, the natural gas savings are greater than the increases in electricity consumption in BTU terms.

The main advantage of the approach proposed here is the fact that it does not require data on where air conditioners are installed. While there are a few surveys that record such data in the US (e.g. RASS, RECS), the spatial coverage is limited and the exact location of the household is often masked for privacy reasons. Our approach circumvents this data limitation, which would be very costly to overcome, by relying on observed electricity consumption from billing data and weather only. The approach outlined here can be adopted for other sectors as well (e.g. health, agriculture).

2. Literature Review

The literature on climate change impacts on energy consumption can be separated into two approaches. In the engineering literature, large-scale bottom-up simulation models are utilized to simulate future electricity demand under varying climate scenarios. The advantage of the simulation model approach is that it allows one to simulate the effects of climate change given a wide variety of technological and policy responses. The drawback to these models is that they contain a large number of response coefficients and make a large number of assumptions about the evolution of the capital stock, for either of which there is little empirical guidance.

The earliest impacts papers adopt this simulation approach and suggest that global warming will significantly increase energy consumption. Cline (1992) provides the earliest study on the impacts of climate change, which relies on an earlier report by the U.S. Environmental Protection Agency (1989). That study of the potential impact of climate change on the United States uses a utility planning model developed by Linder et al. (1987) to simulate the impact on electric utilities in the United States. Baxter and Calandri (1992) focus on California's electricity use. They utilize a partial equilibrium model of the residential, commercial, agriculture, and water pumping sectors to examine total consumption, as well as peak demand. Rosenthal et al. (1995) focus on the impact of global warming on energy expenditures for space heating and cooling in residential and commercial buildings.

The more recent economics literature has frequently applied an econometric approach to impacts estimation, which is the approach we adopt in the current study. While there is a large literature on econometric estimation of electricity demand, the literature on climate change impacts estimation is small and relies on panel estimation of heavily aggregated data or cross-sectional analysis of more micro-level data. The first set of papers attempts to explain variation in a cross section of energy expenditures based on survey data to estimate the impact of climate change on fuel consumption choices. Mansur et al. (2008) and Mendelsohn (2003) endogenize fuel choice, which is usually assumed to be exogenous. They find that warming will result in fuel switching towards electricity. The drawback of the cross-sectional approach is that one cannot econometrically control for unobservable differences across firms and households, which may be correlated with weather/climate. If that is the case, the coefficients on the weather variables and corresponding impacts estimates may be biased. Franco and Sanstad (2008) explain pure time series variation in hourly electricity load at the grid level over the course of a year. They use data reported by the California Independent System Operator for 2004 and regress them on average daily temperature. The estimates show a nonlinear impact of average temperature on electricity load and a linear impact of maximum temperature on peak demand. They link the econometric model to climate model output from three different general circulation models (GCMs) forced using three IPCC scenarios (A1Fi, A2, and B1) to simulate the increase in annual electricity and peak load from 2005-2099. Relative to the 1961-1990 base period, the range of increases in electricity and peak load demands are 0.9 to 20.3 percent and 1.0 to 19.3 percent, respectively. Crowley and Joutz (2003) use a similar approach where they estimate the impact of temperature on electricity load using hourly data in the Pennsylvania, New Jersey, and Maryland Interconnection. Some key differences, however, are that they control for time-fixed effects and define the temperature variable in terms of heating and cooling degree days. They find that a 2°C (3.6°F) increase in temperature results in an increase in energy consumption of 3.8 percent of actual consumption, which is similar to the impact estimated by Baxter and Calandri (1992).

Deschenes and Greenstone (2011) provide the first panel data-based approach to estimating

the impacts of climate change on residential electricity demand. They explain variation in U.S. state-level annual panel data of residential energy consumption using flexible functional forms of daily mean temperatures. The identification strategy behind their paper, which is one we will adopt here as well, relies on random fluctuations in weather to identify climate effects on electricity demand. The model includes state fixed effects, census division by year fixed effects, and controls for precipitation, population, and income. The temperature data enter the model as the number of days in 20 predetermined temperature intervals. The authors find a U-shaped response function where energy consumption is higher on colder and hotter days. The impact of climate change on annual residential energy consumption for the Pacific Census Region (California, Oregon, and Washington) by 2099 is approximately nine percent - yet not statistically different from zero. They are careful to point out that their estimates likely present an upper bound. The panel data approach allows one to control for differences in unobservables across the units of observation, resulting in consistent estimates of the coefficients on temperature.

Aroonruengsawat and Auffhammer (2012) is the first paper using a panel of household level electricity billing data to examine the impact of climate change on residential electricity consumption. They identify the effect of temperature on electricity demand using within-household variation in temperature, which is made possible through variation in start dates and lengths of household billing periods. Since their dataset is a panel, they can control for household fixed effects, month fixed effects, and year fixed effects. Their projected impacts are consistent with the findings by Deschenes and Greenstone, ranging between 1% and 6%.

A separate literature looking at extensive margin adjustments, not necessarily just to higher temperatures, provides significant insight into how households decide to invest in durable goods such as air conditioners and decide how intensively to use them. Biddle (2008) notes that in 1955 the residential air conditioner penetration in the United States was below 2 percent nationally. A quarter of a century later that fraction had risen to 50 percent, with half of those households having installed central air conditioning units. There was significant heterogeneity in the penetration, where half of the residences in the Northeast were air-conditioned, and some urban areas in Texas and

Florida had penetration rates in excess of 90 percent. He argues that on the one hand, the strong downward trend in prices of air conditioners is a potential driver of the increased adoption of this technology. Further, electricity prices dropped significantly during the 1950s and 1960s and then rose again during the 1970s. During this entire period incomes rose substantially, which suggests that the falling costs of installation and operation, combined with rising incomes, drove the adoption of air conditioners during this period. He uses a reduced form econometric model, which accounts for changes in incomes, prices, and weather to explain the heterogeneity in penetration. His estimated income elasticities are positive in most specifications and years, even though they vary widely in size and significance across samples and estimation techniques. The price elasticities are generally negative and often significant, although they vary significantly in magnitude as well.

Sailor and Pavlova (2003) use data on air conditioning penetration for 39 U.S. cities to parameterize a relationship between cooling degree days (CDDs) and market saturation. They show a nonlinear relationship between CDDs and penetration. They further report that a significant number of cities have air conditioning penetration below 80 percent, suggesting that there are two possible margins of adjustment under climate change: Increased adoption of air conditioners and increased usage. Ignoring the adoption decision would lead to an underestimation of future electricity consumption. They estimate a relationship between saturation and CDDs, which uses the notion that current hotter cities have patterns of adoption of cities that are currently cooler, but are expected to be hotter under climate change. The authors show that there is a relationship between the two variables. In their analysis of electricity consumption the authors use state per capita electricity consumption as a proxy for city-level per capita consumption, and their statistical model allows both for increased usage of equipment due to more CDDs and increased adoption of air conditioners due to higher CDDs. They show that simulated increases are significant, and also indicate that taking into account the adoption decision matters in making forecasts. Sailor and Pavlova (2003) note that "Based on these results, Los Angeles' per capita residential electricity consumption is projected to increase by 8% in July for a 20% increase in CDD. If the market saturation were assumed to remain constant, however, the projection would be for only a 5%

increase."

Rapson (2011) estimates a dynamic structural model of air conditioner adoption and simulates the impacts of a carbon tax and energy-efficiency standards. The three relevant findings from his study are that higher energy efficiency of air conditioners raises demand for these units and decreases overall consumption. Second, higher electricity prices have a negative effect on current consumption and a marginal negative effect on adoption. Finally, lower prices of air conditioners lead to higher demand and marginally lower consumption. He uses five cross sections of the Energy Information Administration's Residential Energy Consumption Survey (RECS), which he matches to air conditioner prices and efficiencies.

McNeil and Letschert (2008, 2010) provide a comprehensive model of adoption of air conditioners and appliances using cross-country data. They collect a cross section of data on adoption rates by country from a number of micro-level survey studies and IEA data. They acknowledge that the utility of owning an air conditioner is climate dependent. Further, they point out that the cost of owning an air conditioner is relatively high. In order to break the correlation between income and temperature in their dataset, the authors take a two-step approach. They follow Sailor and Pavlova (2003) and estimate a relationship between saturation and cooling degree days for 39 U.S. cities. They then assume that for a given CDD, developing country air conditioner saturation will approach this frontier, but never exceed it. They assume that diffusion is a function of the climate maximum calculated above and income. They then estimate a relationship between income and air conditioner saturation, which non-surprisingly shows that income is a major driver of the speed of adoption.

Davis and Gertler (2015) provide the most complete estimation of intensive and extensive margin impacts of air conditioning on residential electricity consumption for a large rapidly developing country - Mexico. Like this paper they make use of a massive database on electricity bills to characterize the temperature response on the intensive margin. To characterize their extensive margin impacts, they rely on a large cross sectional survey of appliance ownership across households. They link the two models to simulate impacts of growing income and warming climate on

intensive and extensive margin consumption of electricity. They find significant impacts: For the worst case climate scenario and continued income growth they estimate a 15.4 % increase in electricity consumption by end of century. Once they account for the extensive margin adjustment, the impacts grow to an 83.1% increase in consumption.

This paper adds to this literature in two significant ways. First, it uses a large scale billing dataset for the seventh largest and one of the richest economies in the world - California - to characterize the heterogeneity in temperature response in sample across space. Second, we provide a simple method to characterize both intensive and extensive margin impacts by solely relying on billing data, which makes this a more broadly applicable approach.

3. Data

3.1 Residential Billing Data

The University of California Energy Institute, jointly with California's investor-owned utilities, established a confidential data center, which contains the complete billing history for all households serviced by the four largest investor owned utilities in the state: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), Southern California Gas Company and San Diego Gas and Electric (SDG&E). SDG&E and PG&E are gas and electric utilities, while SoCalGas is gas only and SCE only provides electricity to its customers. Table (1) provides an overview of the data coverage for the four utilities by fuel. We use data for these three utilities for the years 1999-2008 (SCE), 2003-2009 (PG&E) and 2000-2009 (SDG&E).

The dataset contains the complete bill level consumption and expenditure information for the population of single metered residential customers during the decade we have data for. Specifically, we observe an ID number for the physical location (e.g. residence), a service account number (e.g. customer), bill start-date, bill end-date, total electricity consumption (in kilowatt-hours, kWh), and the total amount of the bill (in $) for each billing cycle, as well as the five-digit ZIP code of the premise. Only customers who were individually metered are included in the dataset, hence we cannot

say anything about single metered multi-family homes. We also cannot reliably identify households who have moved and therefore refrain from using this as a source of econometric identification.[1] For the purpose of this paper, a customer is defined as a unique combination of premise and service account number. We can also identify whether a customer receives a low income subsidy on their electricity pricing through a state-level program called CARE. Further, we can also determine which homes are all-electric, meaning that they heat and cool using electricity as well as own electric water heaters.

It is important to note that each billing cycle does not follow the calendar month, and the length of the billing cycle varies across households, with the vast majority of households being billed on a 25-35 day cycle. Hereafter, this dataset is referred to as "billing data." We remove bills with average daily consumption less than 2 kWh from our sample, since we are concerned that these outliers are not regular residential homes, but rather mostly vacant vacation homes. We also remove homes on solar tariffs from our data, since we do not observe total consumption from these homes, but only what they take from the grid, rendering these data useless for the purpose of this exercise.

Our database contains a total of 914 million bills. We observe 658 million bills for "normal" households, which are neither on the CARE tariff nor all-electric homes. In addition we have 92 million bills for all-electric homes in the PG&E and SCE territory. Finally we have 164 million bills for households in all three utility territories on the subsidized CARE tariff. We will treat these three classes of households separately in terms of estimation and simulation.[2]

There is significant variation in bill level consumption across and within households. Because across-household variation may be driven by unobservable characteristics at the household level (e.g., income, physical building characteristics, and installed capital), we will control for unobservable confounders at the household level using fixed effects and use bill-to-bill variation at the household level as our source of identifying variation. To proceed with estimation at the ZIP code level, we

[1]The premise identification number does not change with the occupant of the residence. The service account number, however, changes with the occupant of the residence. This information is subject to sizable measurement error, so we do not make use of "movers" for econometric identification purposes.

[2]The simulations for the CARE and all-electric households did not finish running at the time this paper was sent to UC Irvine.

identify all ZIP codes across the three utilities' territories for which we have at least 500 bills. Figure 1 displays the 1,235 ZIP codes for which we observe such billing data. Our sample covers a large portion of the state and represents approximately 80 percent of California's population.

Since summarizing the consumption data for the 1,235 ZIP codes that ultimately ended up in our estimation cannot be meaningfully presented here, we summarize the consumption, pricing, and weather data for our sample by California Energy Commission climate zones in Table 1. There is great variability in average usage across climate zones, with the highest consumption levels in the hot Central Valley and China Lake (Southeast) zones. The average electricity price paid is similar across zones for our sample, ranging between 13 and 15 cents per kilowatt-hour.

3.2 Weather Data

The daily weather observations to be matched with the household electricity consumption data have been provided by Schlenker & Roberts (2009). This dataset contains daily gridded maximum and minimum temperature for the continental United States at a grid cell resolution of 2.5 miles. Detailed documentation for this dataset are available as part of the supporting material for Schlenker & Roberts (2009). We observe these data for California from 1950 - 2011. In order to match the weather grids to ZIP codes, we have obtained a GIS layer of ZIP codes from the Census Bureau. For small ZIP codes not identified by the Census bureau we have purchased the location of these ZIP codes from a private vendor (zip-codes.com). We matched the Schlenker and Roberts grids to the ZIP code shapes from the census and averaged the daily temperature data across the multiple grids within each ZIP code for each day. For ZIP codes identified as a point, we simply use the daily weather observation in the grid at that point. This leaves us with a complete daily record of minimum and maximum temperature as well as precipitation at the ZIP code level from 1950 - 2010.

3.3 Other Data

We unfortunately only observe bill details about each household and are missing any sociodemographic observables. We do, however, observe the five-digit ZIP code in which each household is located. We purchased socio-demographics at the ZIP code level from a firm aggregating this information from census estimates (zip-codes.com). We only observe these data for a single year (2010). There are 1,752 five digit ZIP codes, which have non-zero population in California. Our sample contains households for 1,235 of these ZIP codes. We do not have sufficient data for households in the remaining 517 ZIP codes. These remaining ZIP codes are either not served by the three utilities, or we do not have a sufficient number of bills (500) for them. Table 2 shows summary statistics for both the ZIP codes in our sample and the ZIP codes for which we do not have electricity data. The ZIP codes in our sample represent 77.09 percent of California's population. The ZIP codes in our sample are more populated, younger, richer, have more expensive homes, have slightly more persons per household, and have a lower proportion of Caucasians and a higher proportion of Hispanics and Asians. There seems to be no statistically significant difference in summer climate, which is important for our purposes. Taking these differences into consideration is again important when judging the external validity of our estimation and simulation results. We will not make explicit use of this information in our first-stage regression, but control for the observable sources of variation in our cross-sectional second stage, which does not allow for a fixed effects strategy by design. The variables we will make use of in the second stage are income and summer climate.

4. Econometric Estimation Strategy

4.1 Intensive Margin: The Usage Response to Temperature

Equation 1 below shows our main estimating equation, which is a simple log-linear equation estinated separately for each ZIP code j. This estimating equation has been commonly employed in aggregate electricity demand estimation and climate change impacts estimation (e.g., Deschenes

and Greenstone 2011, Davis and Gertler, 2015).

$$\log(q_{it}) = \sum_{p=1}^{14} \beta_j^p D_{pit} + \gamma Z_{it} + \alpha_i + \phi_m + \psi_y + \varepsilon_{it} \quad (1)$$

where $\log(q_{it})$ is the natural logarithm of household i's daily average electricity consumed in kilowatt-hours during billing period t. D_{pit} are our measures of temperature, which we discuss in detail below. Z_{it} are observed confounders at the household level, α_i are time invariant household fixed effects, ϕ_m are month of year fixed effects, and ψ_y are year fixed effects. ε_{it} is a stochastic error term. As bills do not overlap with calendar months and years perfectly, ϕ_m and ψ_y are assigned as shares to individual bills according to the share of days in a bill for each month and year.

For estimation purposes, our unit of observation is a unique combination of premise and service account number, which is associated with an individual and structure. We thereby avoided the issue of having individuals moving to different structures with more or less efficient electricity consuming capital or residents with different preferences over electricity consumption moving in and out of a given structure.

California's housing stock varies greatly in its energy efficiency and installed energy-consuming capital. Further, California's population is not randomly distributed across ZIP codes. We suspect that there may be differences in attitudes towards cooling, installed capital, quality of construction across ZIP codes, and the associated demographics and capital. The key novelty in this paper is that we estimate equation (1) *separately* for each of the 1,235 ZIP codes displayed in Figure 1. This is possible, since we observe such a large amount of data. While big data in past was a problem overcome by sampling, it provides an opportunity in this context: the identification of 1,235 electricity temperature response functions across space. The motivation for doing this is that we would expect the relationship between consumption and temperature to vary across these ZIP codes according to the penetration of air conditioners and the resident population's propensity to use these.

The main variables of interest in this paper are those measuring temperature. Following recent trends in the literature we include our temperature variables in a way that imposes a minimal

number of functional form restrictions in order to capture potentially important nonlinearities of the outcome of interest - electricity consumption - in weather (e.g., Schlenker and Roberts 2006; Deschenes and Greenstone 2011, Davis and Gertler (2015)). We achieve this by sorting each day's mean temperature experienced by household i into one of 14 temperature bins. For the purposes of this study, we use the same set of bins for each ZIP code in the state. In order to define a set of temperature bins we split the state's temperature distribution into a set of percentiles and use those as the bins sorting. Aroonruengsawat and Auffhammer (2012) show that the alternative approach of using equidistant five-degree bins yields almost identical results. We therefore refrain from following this approach, due to the computational intensity involved. As a result, not each ZIP code will have observations in each bin. The northern ZIP codes, for example, do not experience days in the upper bins, while the southwestern parts of California have few days in the coldest bins.

We split the temperature distribution into deciles, yet break down the upper and bottom decile further to include buckets for the first, fifth, ninety-fifth, and ninety-ninth percentile to account for extreme cold/ heat days. We therefore have a set of 14 buckets which we use for each household, independent of in which climate zone the household is located.[3] For each household and bill, we count the number of days the mean daily temperature falls into each bin and record this as D_{pit}. The main coefficients of interest are the fourteen β_j^p coefficients, which measure the impact of one more day with a mean temperature falling into bin p on the log of household daily electricity consumption in ZIP code j. For small values, β_j^p's interpretation is approximately the percent increase in household electricity consumption during a billing period from experiencing one additional day in that temperature bin.

The last four columns in table (1) display the heterogeneity in the weather distribution across climate zones for the state of California. The Central Valley and the southeastern areas of the state for example, are significantly hotter on average, as shown by the larger share of days spent above 86° Fahrenheit. Figure 2 displays the daily average temperature for the months of June, July and August averaged over the years 1950-2010. This is a reasonable measure of summer climate (a sixty

[3]The cutoffs for the bins are 32, 42, 46, 51, 54.5, 58, 60, 63, 67, 70, 75, 79 and 86 degrees Fahrenheit mean daily temperature.

year average instead of the usual thirty year average). Figure (2) shows that the Central Valley non-coastal areas of Southern California are very warm during the summer months. We would expect for these areas to have a significantly more temperature sensitive electricity consumption response than the cooler coastal areas of Northern California and higher altitude settings in the Sierras.

Z_{it} is a vector of observable confounding variables, which vary across billing periods and households. The first of two major confounders that we observe at the household level are the average electricity price for each household for a given billing period. California utilities price residential electricity on a block rate structure. The average price experienced by each household in a given period is therefore not exogenous, since marginal price depends on consumption (q_{it}). Identifying the price elasticity of demand in this setting is problematic, and a variety of approaches have been proposed (e.g., Hanemann 1984; Reiss and White 2005; Ito, 2014). The maximum likelihood approaches are computationally intensive and given our sample size cannot be feasibly implemented here. We have run our models by including average price directly, instrumenting for it using lagged prices and omitting it from estimation. The estimation results are almost identical for all three approaches, which is reassuring. While one could tell a story that higher temperatures lead to higher consumption, and therefore higher marginal prices for some households, this shift in parameter estimates seems to be negligible given our estimation results. In the estimation and simulation results presented in this paper, we omit the average price from our main regression. The second major time-varying confounder is precipitation in the form of rainfall. We control for rainfall using a second-order polynomial in all regressions.

To credibly identify the effects of temperature on the log of electricity consumption, we require that the residuals conditional on all right-hand side variables be orthogonal to the temperature variables, which can be expressed as $E[\varepsilon_{it} D_{pit} | D_{-pit}, Z_{it}, \alpha_i, \phi_m, \psi_y] = 0$. Since we control for household fixed effects, identification comes from within household variation in daily temperature after controlling for confounders common to all households (e.g., business cycle effects) and rainfall. We estimate equation (1) for each of the 1,235 ZIP codes in our sample using a least-squares fitting

criterion and a household level clustered variance covariance matrix. This approach serves as the first stage in our overall methodology and serve as the basis for our estimates of intensive margin adjustment due to climate change. We must make the assumption that response to slowly changing climate over this four-year period is small in order to be able to interpret our coefficients as the intensive margin adjustment to the changes in usage of existing equipment in response to changing temperature, which we deem to be reasonable.

4.2 The Long-Run Response to Temperature

In a warmer world, existing air conditioners will be run for more hours, which we call the intensive margin adjustment. The second margin of adaptation is the installation of additional air conditioners in existing homes and new construction. One can easily imagine that if San Francisco's future climate resembles that of current day Fresno during the summers, the wealthy residents of San Francisco will install (additional) cooling equipment in their homes. We provide an attempt to quantify the magnitude of this response. We estimate equations of the following form:

$$\beta_j^p = \delta_1 + \delta_2 C_j^p + \boldsymbol{\delta_3 Z_j} + \eta_j \tag{2}$$

where β_j^p is a measure of ZIP code j's temperature responsiveness in bin $p \in [10; 14]$ as estimated in equation (1). We would expect there to be a response only in the upper portion of the temperature response curve, where cooling occurs. A common threshold for the uptick in the temperature response curve, which we will show is valid for our data in the results section, is 65 degrees Fahrenheit. The bin containing this temperature is our omitted bin in the estimation, hence we see if hotter ZIP codes have a steeper temperature response curve for bins above 65 degrees.

The variable C_j^p in equation (2) is the share of days ZIP code j experienced in temperature bin p during the sample years 2000-2010 from our ZIP code level weather data produced from the Schlenker and Roberts (2009) data. C_j^p is bounded by 0 and 1 and adds to one if one were to sum it across all temperature bins from 1-14. The variable(s) Z_j are any confounders that may affect

the temperature response of the population in ZIP code j. The main confounder we consider here is income, as higher-income households are thought to more easily afford the capital expenditure of an air conditioner and its associated operating expense (Rapson 2011). While we will not use the estimated coefficient on income in our simulation later, controlling for it ensures that we do not confound the temperature extensive margin adjustment by income. If individuals sort into climate according to income, failing to control for these factors would bias our estimated climate response. In terms of estimation, we could estimate five separate equations of type (2) or estimate a pooled regression allowing for flexibility in the δ_2 coefficient for higher bins. As discussed below we chose to estimate a pooled model, which restricts the coefficient on income to be identical for all bins, yet controls for bin dummies. This provided more stable estimation results than the separate equations.

Finally, estimating equation (2) via Ordinary Least Squares is not optimal, as the dependent variables are estimated coefficients. As we observe the estimated standard error for each of the almost 6,000 β_j^p, we run variance weighted regressions for model (2).

5. Estimation Results

5.1 The Usage Response to Temperature

As discussed in the previous section, we estimate equation (1) for each of the 1,235 ZIP codes that have more than 500 bills. While we cannot feasibly present the 1,235 estimated temperature response functions (which are comprised of 13 parameter estimates each), we can display the distribution of the temperature response curves in a fan plot, which is shown in figure (3). The thick black line in figure (3) displays the median temperature response curve across the 1,235 ZIP codes. As the regression has average daily consumption on the left hand side and the number of days out of a normalized 30 spent in each bin on the right hand side, the coefficients indicate the percent change in average daily consumption from one additional day spent in a given bin relative to a day in the 65 degree bin. The curve has the expected U-shape with a steep positive slope above 65 degrees and a shallower negative slope at temperatures below 65 degrees. The bottom of the

U-shaped response curve is right near the omitted bin of 65 degrees. Examining the distribution of the slopes of the temperature response across the state displays patters similar to those found by Aroonruengsawat and Auffhammer (2012). They demonstrated that within a given temperature bin, there is significant variation in temperature response across the state - depending on physical location. They only estimate sixteen distinct temperature response curves, which makes it difficult to examine the source of variation in slope at higher temperatures, which is what we will do in the next section. Figure 3 displays the significant heterogeneity in temperature response via the shaded fan areas. The palest grey fan indicates bounds the 5^{th} to 95^{th} percentile of the distribution. Each darker shade of grey increments the interval by 10%. What we see here is that there are significant numbers of ZIP codes with an extremely steep temperature response as well as a significant number of ZIP codes with an almost flat temperature response. We have produced analogous pictures for the CARE households and the all-electric households. These graphs are in the supplemental materials to this paper. The CARE distribution has a slightly shallower slope at both high and low temperatures. The all-electric distribution has slightly steeper slopes at higher and lower temperatures, which is to be expected as these houses tend to be older and heating and cooling are conducted using electricity, not natural gas.

5.2 The Long-Run Response to Temperature

As discussed in Section 4.2, we exploit the 1,235 estimated temperature response curves and examine whether we can explain variation in temperature response through cross-sectional variation in "climate" as well as income. The left hand side variable is our measure of temperature response of electricity consumption for each of the bins 10-14, which we estimated for each ZIP code i in the previous step using equation (1). On the right hand side we control for the percent of days spent in the respective bins and income. We run a pooled regression, the results of which are shown in table (3). As the dependent variable is an estimated coefficient, we employ variance weighted regressions here. Model (1) simply regressed the slope for the 5,612 ZIP/bin observations on the share of days spent in the bin. The coefficient is negative, which speaks to omitted variables issues. In model (2)

we control for income and somewhat surprisingly the coefficient does not move. Model (3) controls for bin fixed effects, which allows for separate intercepts for each bin and the coefficient carries the correct positive sign, which indicates that having more days in the bins, which usually see some cooling, results in a steeper temperature response. Model (4) allows for a differential shift in the temperature response function for the three highest bins and as expected the shift is significant and much larger than the pooled estimate. We use these estimated coefficients from model (4) as the basis for our simulation.

It is important to note that individuals do not randomly select into locations. There is evidence from a large literature on economic sorting models that individuals self-sort into locations that match their preferred amenities, given their income level. The regressions in table (3) therefore carry no causal interpretation, but are merely correlations. We would expect experiencing more time in hotter temperatures would carry a positive coefficient for the higher temperature bins, which it does.

We use the results from model (4) in table (3) to simulate the impacts of climate change on the slope of each ZIP code's temperature response curve. In the next section, we will generate a large number of counterfactual climate futures from 18 General Circulation Models (the technical term for climate models) and two different scenarios of emissions. We will use these ZIP code level climate futures to shift each ZIP code's temperature response curve for bins 10-14 using the estimates from table (3). This simulated shift in the temperature response curve will allows to quantify the extensive margin adaptation response.

6. Electricity Consumption Increases Due to Climate and Population Change

In this section, we simulate the impacts of climate change on electricity consumption under two different emissions scenarios using eighteen different climate models from the latest round of the IPCC assessments (AR5, CMIP5) in their downscaled form. We conduct three different simulations.

The first simulation holds population growth constant and only simulates electricity consumption per household using the first-stage estimates, which do not allow for changes in the extensive margin. In a second simulation we incorporate the extensive margin adjustments conducted in sections 4.2 and 5.2. In a final simulation we allow for population growth. For each simulation we can calculate the trajectory of aggregate electricity consumption from the residential sector until the year 2099, which is standard in the climate change literature. We provide simulated impacts for the periods 2020-2039, 2040-2059, 2060-2079 and 2080-2099.

6.1 Temperature Simulations

The simulation for this section uses the climate response parameters estimated in Section 5.1. Using these estimates as the basis of our simulation has several strong implications. Using the only the first stage parameters via equation (1) implies that the climate responsiveness of consumption within climate zones remains constant throughout the century.

As is standard in this literature, the counterfactual climate is generated by a general circulation model (GCM). These numerical simulation models generate predictions of past and future climate under different scenarios of atmospheric greenhouse gas (GHG) concentrations. The quantitative projections of global climate change conducted under the auspices of the IPCC's fifth assessment report (AR5) and applied in this study are driven by modeled simulations of two sets of projections of twenty-first century social and economic development around the world, the so-called "RCP4.5" and "RCP8.5" storylines. The number after the RCP stands for the likely increase in forcing from the scenario by end of century relative to preindustrial values in Watts per square meter. In terms more familiar to most economists, RCP4.5 is expected to result in a warming of 1.8 °C with a likely range of 1.1. to 2.6 °C. This is a very optimistic scenario as attaining a goal of warming less than 2 degrees is unlikely. RCP 8.5 is the worst case scenario and expected to result in a warming of 3.7 °C with a likely range of 2.6 to 4.8 °C.

We simulate consumption for each scenario using the 18 downscaled GCMs from the IPCC's CMIP5 database. The downscaled temperature scenarios were drawn from a statistical downscaling

of global climate model (GCM) data from the Coupled Model Intercomparison Project 5 (CMIP5, Taylor et al. 2010) utilizing a modification (Hegewisch, Abatzoglou, in prep.) of the Multivariate Adaptive Constructed Analogs (MACA, Abatzoglou and Brown, 2012) method with the Livneh (Livneh et.al.,2013) observational dataset as training data. These were provided to us by the MACA project at the University of Idaho. We matched the fine scale grids of the downscaled climate data to ZIP codes in the same fashion we matched the Schlenker and Roberts (2009) weather grids. We calculated future climate by adding the predicted change in monthly temperature for each model, scenario and period to our baseline weather data to avoid local biases as the MACA project does not use the same weather data as its training data set.[4]

To obtain estimates for a percent increase in electricity consumption for the representative household in ZIP code j and period $t+h$, we use the following relation:

$$\frac{q_{j,t+h}}{q_{j,t}} = \frac{exp(\sum_{p=1}^{14} b_{pj} D_{pj,t+h})}{exp(\sum_{p=1}^{14} b_{pj} D_{pj,t})} \tag{3}$$

The top panel of figure (4) displays the projected increases in household residential electricity consumption across the 1,235 ZIP codes for each of the four projection periods and the 18 GCMS for RCP8.5 (RCP4.5 is in the supplementary materials). The box plots display tremendous variation across time (the box and whiskers plots for each model are shown in increasing temporal order for each model), across models and within models. It is quite clear that median impacts are increasing over time and impacts range from the negative teens to increases approaching 50% for some ZIP codes. While this figure is useful in displaying the variability in projections, it does not display the spatial variability in intensive margin impacts for the average household across ZIP codes. Figure (5) tries to convey this by plotting the predicted impact for the average household by end of century using the ensemble average prediction across all 18 GCMs for RCP8.5. What this graph shows is that the ZIP codes in the Central Valley and non-coastal Southern California are projected to experience the largest increases in household electricity consumption. This is due to the combination of the slope of the temperature response function and projected warming from the GCMs. These

[4]A detailed description of the climate model output is available at http://maca.northwestknowledge.net/.

projections ignore potential extensive margin impacts, which we turn to next.

For each ZIP code, climate model and scenario, we calculate the temperature response curve using model (4) in table (3). As the temperature distribution shifts to the right for the vast majority of ZIP codes in California, a higher share of days in the higher bins is projected under both climate change scenarios for most models. It is impractical to show the almost 44,460 counterfactual response curves, yet figure (6) collapses the temperature response curves across ZIP codes by projection period. The top panel displays the population weighted state wide response curve in-sample in black and the projected future response curves in blue and red. As expected the response curve tilts up more and more over time. The bottom panel repeats this exercise for RCP 4.5, which results in significantly less movement. We now use the extensive margin adjusted response functions to simulate impacts of climate change on electricity consumption. The bottom panel of figure (4) displays the same box and whisker plots as we did for the intensive margin simulations earlier. What stands out from this graph is an almost uniform upward shift in the medians across models and increased variability across models - especially at the high end. Figure (7) displays the impacts on the average household in a ZIP code using the ensemble average of GCMs and RCP 8.5 by end of century across the state. Compared to the impacts shown in figure (5), the map is significantly darker, indicating a noticeable increase in consumption across the state. To make this more apparent, figure (8) displays the difference between the extensive and intensive margin estimates. The figure shows that these extensive margin impacts will be felt most strongly in the Central Valley and non-coastal areas of Southern California.

While these maps are instructive, it is hard to determine how big the overall impact of allowing for extensive margin adjustment is. Table 4 therefore shows the overall population-weighted increases in total electricity consumption averaged across the 18 climate models and the two RCPs - with and without extensive margin adjustments. The first thing to notice from this table is that accounting for the extensive margin adjustments results in a significant difference in simulated impacts, which is consistent with the findings in Davis and Gertler (2015) for Mexico. For RCP4.5 by the end of the century, accounting for extensive margin impacts almost triples the estimated

impacts. For RCP8.5 failing to account for extensive margin impacts more than triples the estimated impacts. The second noteworthy fact is that even until 2059, the estimated impacts are relatively small - strictly less than 5% even for the worst case scenario. In terms of electricity planners planning horizon this is essentially infinitely far into the future and the magnitude of the impacts is in the noise. By the end of the century, however the impacts are larger, yet their magnitudes are small enough that not overly optimistic assumptions about technological change related to energy efficiency should more than be able to offset these gains.

6.2 Temperature and Population Simulations

California has experienced an almost seven-fold increase in its population since 1929 (BEA 2008), and California's population growth rate over that period (2.45 percent) was more than double that of the national average (1.17 percent). Over the past 50 years California's population has grown by 22 million people to almost 37 million in 2007 (BEA 2008). To predict what the trajectory of California's population will look like until the year 2100, many factors have to be taken into account. The four key components driving future population are net international migration, net domestic migration, mortality rates, and fertility rates. The State of California provides forecasts 55 years into the future, which is problematic, since we are interested in simulating end-of-century electricity consumption. The Public Policy Institute of California has generated a set of population projections until 2100 at the county level, and we obtained these from Sanstad et al. (2009).

The three sets of projections developed for California and its counties are designed to provide a subjective assessment of the uncertainty of the state's future population. The projections present three very different demographic futures. In the low series, population growth slows as birth rates decline, migration out of the state accelerates, and mortality rates show little improvement. In the high series, population growth accelerates as birth rates increase, migration increases, and mortality declines. The middle series, consistent with (but not identical to) the California Department of Finance projections, assumes future growth in California will be similar to patterns observed over the state's recent history - patterns that include a moderation of previous growth rates but still large

absolute changes in the state's population. In the middle series, international migration flows to California remain strong to mid-century and then subside, net domestic migration remains negative but of small magnitude, fertility levels (as measured by total fertility rates) decline slightly, and age-specific mortality rates continue to improve.

The high projection is equivalent to an overall growth rate of 1.47 percent per year and results in a quadrupling of population to 148 million by the end of the century. The middle series results in a 0.88 percent annual growth rate and 2.3-fold increase in total population. The low series is equivalent to a 0.18 percent growth rate and results in a population 18 percent higher than today's. Projections are available at the county level and not at the ZIP code level. We therefore assume that each ZIP code in the same county experiences an identical growth rate.

Table 5 displays the simulated aggregate changes in electricity consumption all three population growth scenarios under the two scenarios of climate change averaged across the 18 GCMs using the full intensive and extensive margin adjustment. It is not surprising to see that population growth has much larger consequences for simulated total electricity consumption compared to climate uncertainty or price uncertainty. The simulations for the low forcing scenario RCP4.5 and the low population growth scenario show a 28 percent increase in consumption, which is largely due to projected increases in population. For the RCP8.5 scenario and the high population growth scenario, the predicted increases are pushing a tripling in consumption. This, unsurprisingly, stresses that population trajectories are much bigger drivers of residential electricity demand than climate change.

7. Conclusions

In the residential sector, one of the most widely discussed modes of adaptation to higher temperatures due to climate change is the increased demand for cooling and decreased demand for heating in the built environment. Due to its mild climate and heavy reliance on natural gas, California's residential sector uses relatively little electricity for heating. It is therefore expected that the de-

mand for electricity will increase as households operate existing air conditioners more frequently, and in many regions will install air conditioners where there currently are few. This paper provides reduced form estimates of changes in electricity consumption due to increased use of installed cooling equipment under a hotter climate. This study adds to the literature by incorporating the change in temperature responsiveness due to likely increases in air conditioner penetration under climate change using a two-stage method. The advantage of the proposed methods lie in its relative simplicity and the fact that it only requires data on electricity consumption and not on installed cooling equipment. The paper shows that accounting for extensive margin adjustments will lead to statistically and economically significantly higher projections of electricity consumption.

It is important to note that we do not and cannot model changes in electricity consumption due to improvements in the efficiency of heating and cooling equipment and/or buildings. These effects will offset some or all of the gains in consumption outlined in this paper. Further, the extensive margin adjustments in this paper cannot meaningfully control for changes in urban form, urban heat island effects, or other variables potentially leading to a higher response, which may be correlated with temperature. We leave the study of these effects to future work.

References

[1] Aroonruengsawat, Anin, and Maximilian Auffhammer. 2012. Correction: Impacts of Climate Change on Residential Electricity Consumption: Evidence From Billing Data. California Energy Commission White Paper.

[2] Bajari, Patrick L., Jane Corinne Cooley, Kyoo-il Kim, and Christopher Timmins. Forthcoming. A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution. American Economic Review.

[3] Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, Joseph S. Shapiro. 2016. "Adapting to Climate Change: The Remarkable Decline in the U.S. Temperature-Mortality Relationship over the 20th Century." Journal of Political Economy. 124(1): 105-159

[4] Baxter, L. W., and K. Calandri. 1992. "Global warming and electricity demand: A study of California." Energy Policy 20(3): 233-244.

[5] Biddle, Jeff. 2008. "Explaining the spread of residential air conditioning, 1955-1980." Explorations in Economic History 45: 402-423.

[6] BEA (Bureau of Economic Analysis). 2008. Regional Economic Accounts. Washington, D.C. www.bea.gov/regional/spi/default.cfm?series=summary.

[7] Cline, W. R. 1992. The economics of global warming. Washington: Institute for International Economics.

[8] Crowley, C., and F. Joutz. 2003. Hourly electricity loads: Temperature elasticities and climate change. In 23rd U.S. Association of Energy Economics North American Conference.

[9] Davis, Lucas W. and Paul J. Gertler. 2015 "Contribution of air conditioning adoption to future energy use under global warming" PNAS doi:10.1073/pnas.1423558112

[10] Deschenes, O., and M. Greenstone. 2011. "Climate Change, Mortality and Adaptation: Evidence from annual fluctuations in weather in the U.S." American Economic Journal: Applied Economics. 3: 152-185.

[11] Energy Information Administration (EIA). 2008. State Energy Data System. Washington, D.C. http://www.eia.doe.gov/emeu/states/_seds.html.

[12] Energy Information Administration. 2011. "Air conditioning in nearly 100 million U.S. homes." http://www.eia.gov/consumption/residential/reports/air_conditioning09.cfm?src=email.

[13] Franco, G., and A. Sanstad. 2008. "Climate change and electricity demand in California." Climatic Change 87:139-151.

[14] Hayhoe, K., M. Robson, J. Rogula, M. Auffhammer, N. L. Miller, J. VanDorn, and D. Wuebbles. 2010. "An integrated framework for quantifying and valuing climate change impacts on urban energy and infrastructure: A Chicago case study." Journal of Great Lakes Research. 36: 94-105.

[15] Hanemann, W. 1984. "Discrete/Continuous Models of Consumer Demand." Econometrica 52: 541-561.

[16] Houthakker, H. S., and L. D. Taylor. 1970. Consumer Demand in the United States: Analyses and Projections. Cambridge, Mass: Harvard University Press.

[17] Intergovernmental Panel on Climate Change (IPCC). 2000. Emissions Scenarios. Cambridge, UK: Cambridge University Press.

[18] Intergovernmental Panel on Climate Change (IPCC). 2013. Climate Change 2013: The physical science basis. Cambridge, UK: Cambridge University Press.

[19] Ito, Koichiro. 2014. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing". American Economic Review, 104(2): 537-63.

[20] Linder, P., M. Gibbs, and M. Inglis. 1987. Potential impacts of climate change on electric utilities. Report 88-2. Albany, New York: New York State Energy Research and Development Authority.

[21] Mansur, E., R. Mendelsohn, and W. Morrison. 2008. "Climate change adaptation: A study of fuel choice and consumption in the U.S. energy sector." Journal of Environmental Economics and Management 55(2): 175-193.

[22] Maurer, E. P., and H. G. Hidalgo. 2008. "Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods." Hydrology and Earth System Sciences 12:551-563.

[23] McNeil, M., and V. Letschert. 2008. Future Air Conditioning Energy Consumption in Developing Countries and what can be done about it: The Potential of Efficiency in the Residential Sector. LBNL Working Paper. http://escholarship.org/uc/item/64f9r6wr.

[24] McNeil, M., and V. Letschert. 2010. "Modeling diffusion of electrical appliances in the residential sector." Energy and Buildings 42: 783-790.

[25] Mendelsohn, R. 2003. The Impact of Climate Change on Energy Expenditures in California. Appendix XI in Wilson, Tom, and Larry Williams, Joel Smith, Robert Mendelsohn. Global Climate Change and California: Potential Implications for Ecosystems, Health, and the Economy. Consultant report 500-03-058CF to the Public Interest Energy Research Program, California Energy Commission.

[26] Miller, N. L., K. Hayhoe, J. Jin, and M. Auffhammer. 2008. "Climate change, extreme heat, and electricity demand." J. Applied Meteorology and Climatology 47(6): 1834-1844.

[27] Rapson. D. 2011. "Durable Goods and Long-Run Electricity Demand: A Case Study of Air Conditioner Purchase Behavior." Working Paper, UC Davis.

[28] Reiss, P. C., and M. W. White. 2005. "Household Electricity Demand Revisited." The Review of Economic Studies. 72:853-883.

[29] Rosenthal, D., H. Gruenspecht, and E. Moran. 1995. "Effects of global warming on energy use for space heating and cooling in the United States." The Energy Journal 16:77-96.

[30] Sailor, D., and A. Pavlova. 2003. "Air conditioning market saturation and long-term response of residential cooling energy demand to climate change." Energy 28 (9): 941-51.

[31] Sanstad, A., H. Johnson, N. Goldstein, and G. Franco. 2009. Long-Run Socioeconomic and Demographic Scenarios for California. PIER Report CEC-500-2009-013-D.

[32] Schlenker, W., and M. Roberts. 2006. "Nonlinear Effects of Weather on Corn Yields." Review of Agricultural Economics 28(3): 391-398.

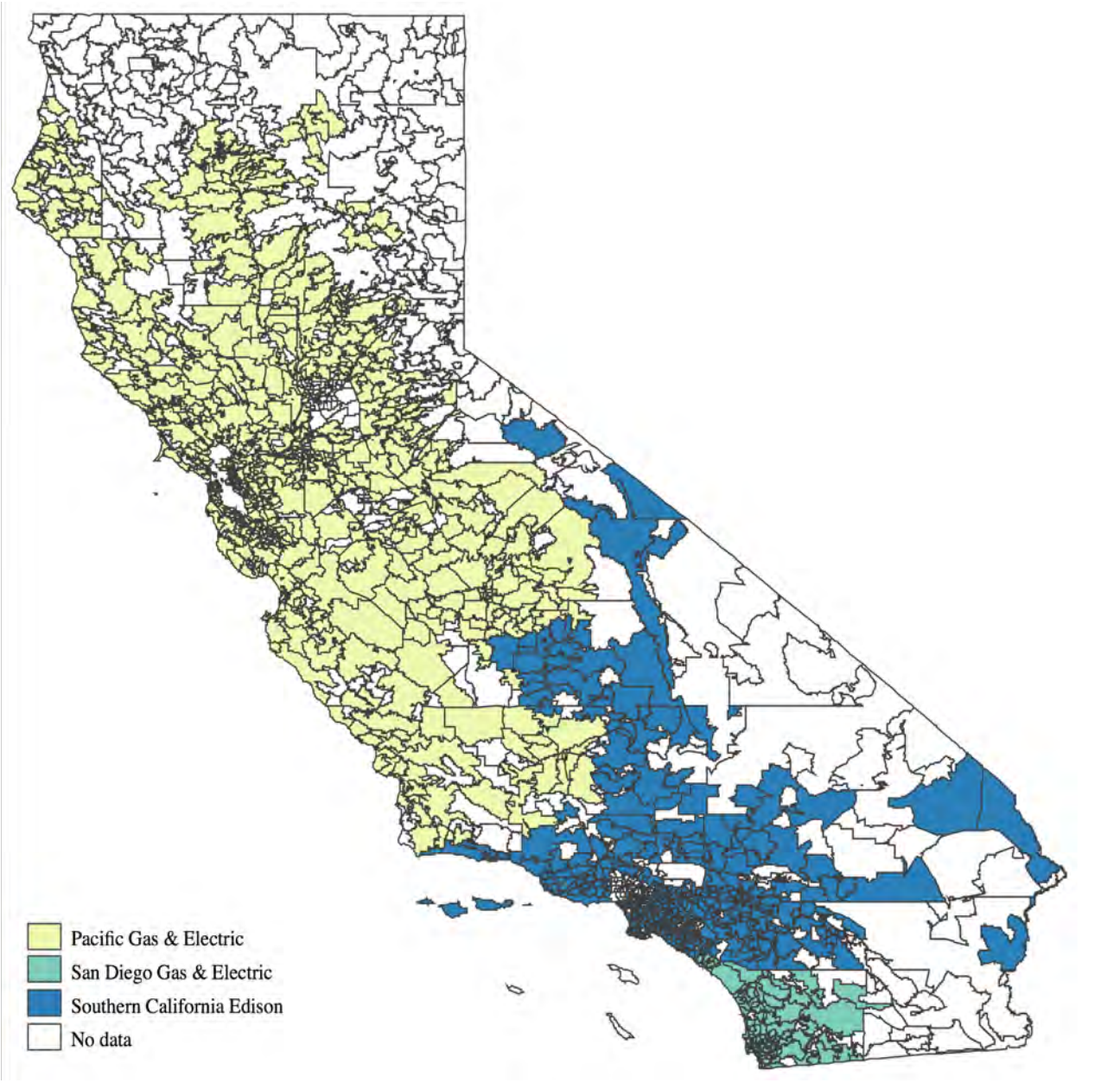
[33] Schlenker and Roberts. 2009. "Nonlinear Temperature Effects indicate Severe Damages to U.S. Crop Yields under Climate Change." Proceedings of the National Academy of Sciences, 106(37), September 15 2009, p.15594-15598.

[34] U.S. Environmental Protection Agency. 1989. The potential effects of global climate change on the United States. Joel B. Smith and Dennis Tirpak, eds. Washington: Environmental Protection Agency.

Table 1: Electricity and Natural Gas Bills by Utility

Utility	Electricity Years	# of Bills	Gas Years	# of Bills
PG&E	2003-2009	342 Million	2004-2014	587 Million
SDG&E	2000-2009	153 Million	2008-2015	74 Million
SCE	1999-2008	469 Million		
SoCalGas			2010-2015	267 Million
Total		964 Million		928 Million

Figure 1: ZIP Codes with observed residential electricity bills by Investor Owned Utility.



Notes: The map above displays the five-digit ZIP codes for which we have more than 500 bills over the estimation period from either PG&E, SCE or SDG&E. Zip codes with no data either have fewer than 500 bills total or a reserved by one of California's many municipal utilities.

Figure 2: California's Summer Climate: Average Daily Temperature July - August 1950-2010

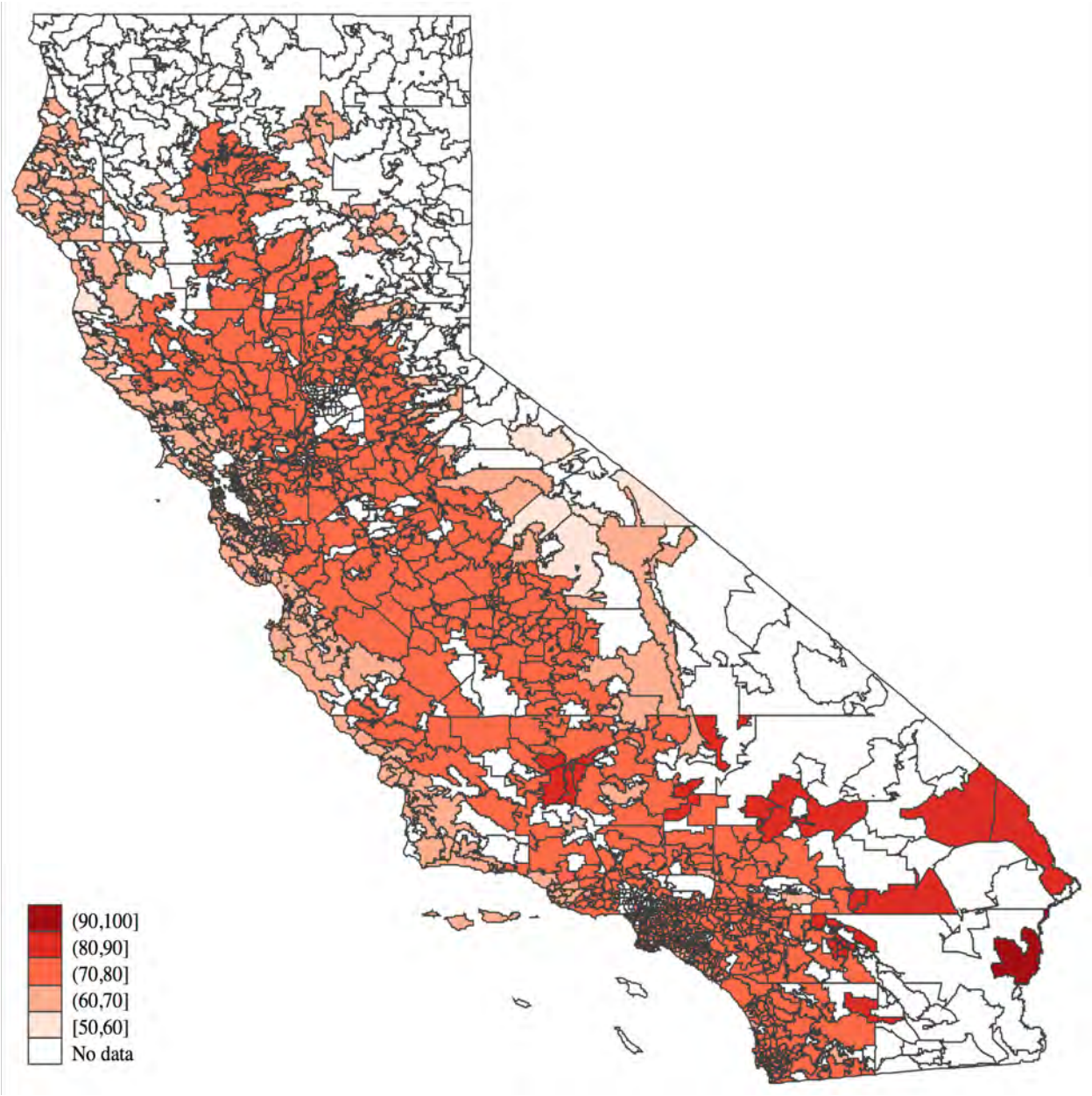
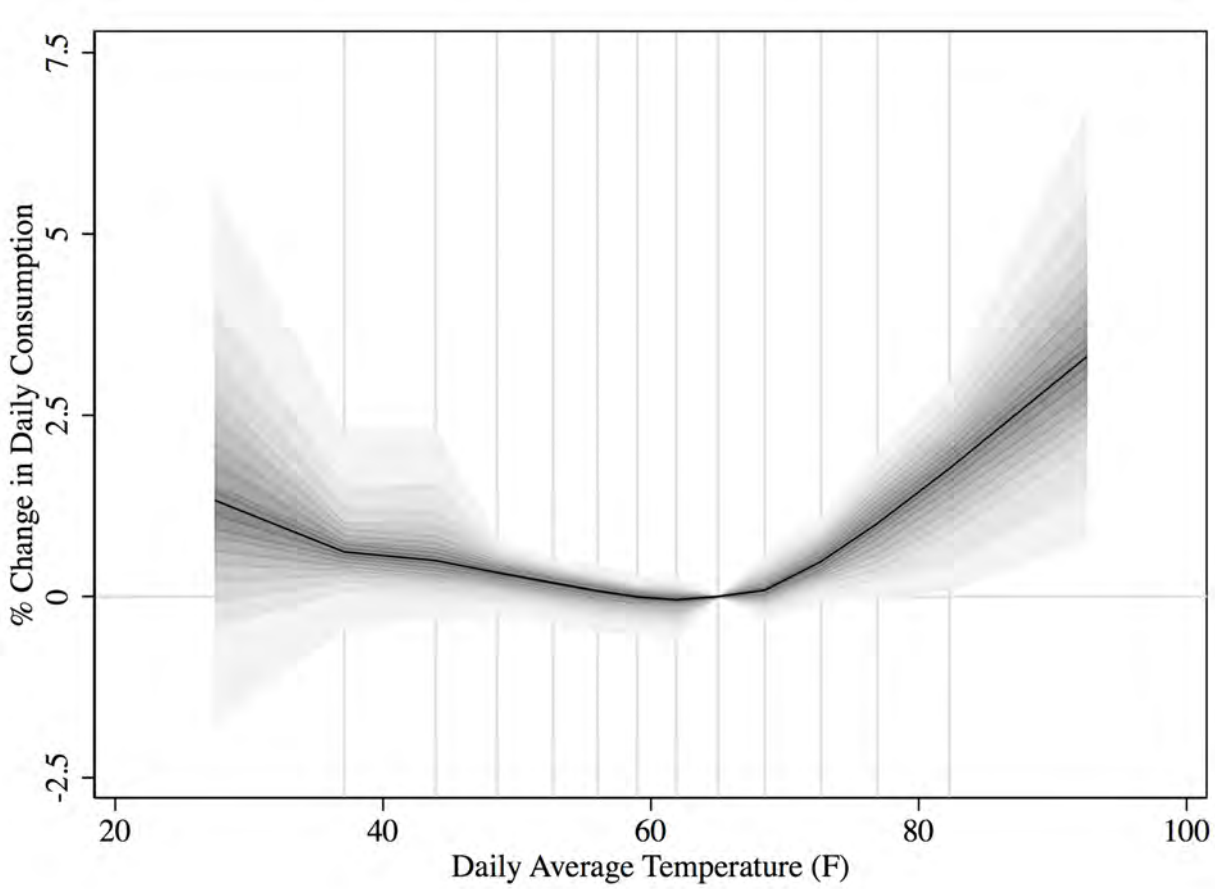


Figure 3: Distribution of Estimated Temperature Response Coefficients across ZIP Codes



Notes: This figure displays the empirical distribution of the estimated temperature response function across ZIP codes in the sample across percentile temperature bins. The lightest gray shading indicates the range of the 5^{th} to 95^{th} percentile. Each darker shading represents a 5% increase in the percentile. The solid black line represents the median temperature responsiveness. The vertical grey lines indicate the cutoffs of the temperature bins.

Figure 4: Intensive margin [Top Panel] and Extensive Margin [Bottom Panel] per household impacts across ZIP codes and Climate Models.

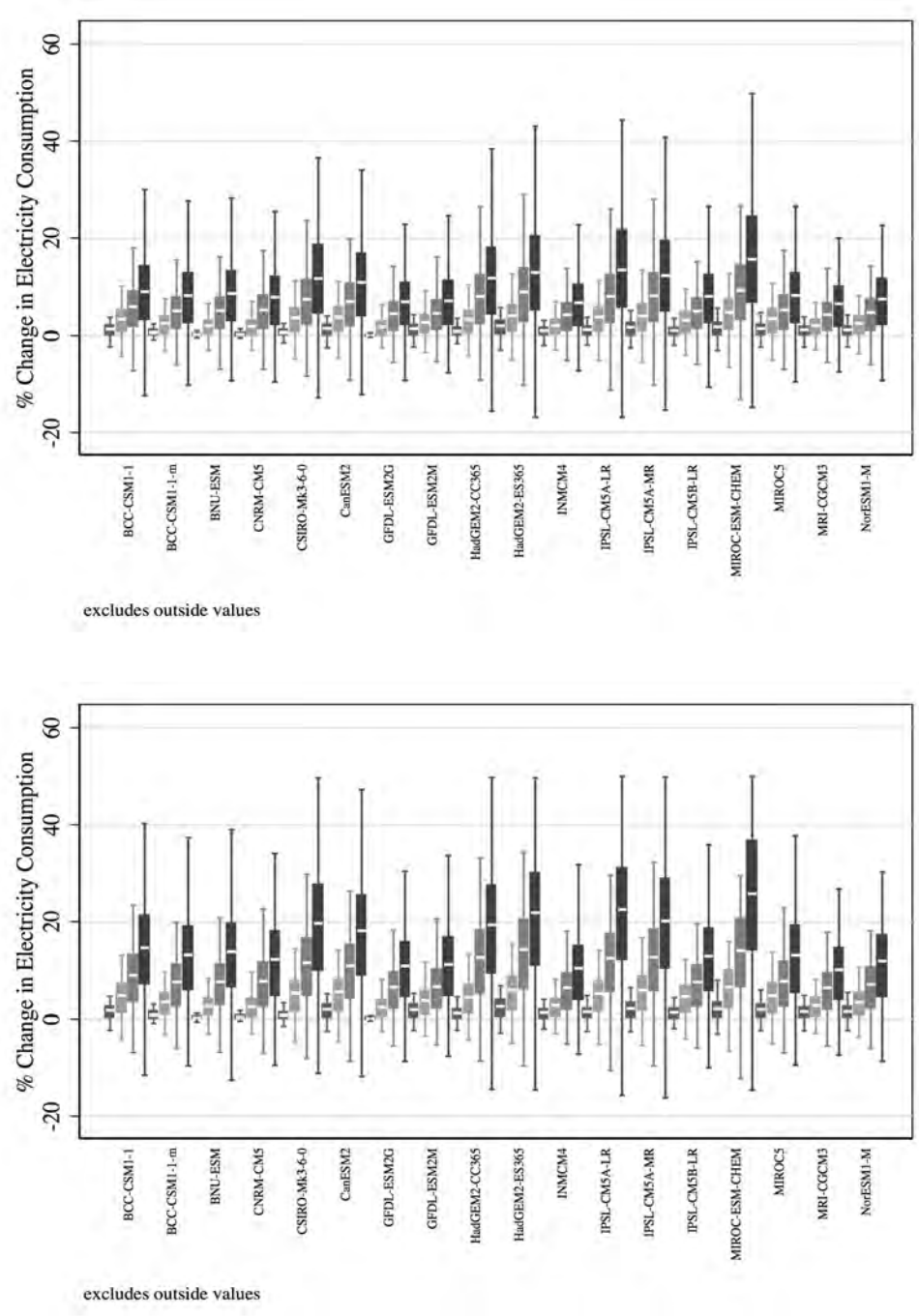
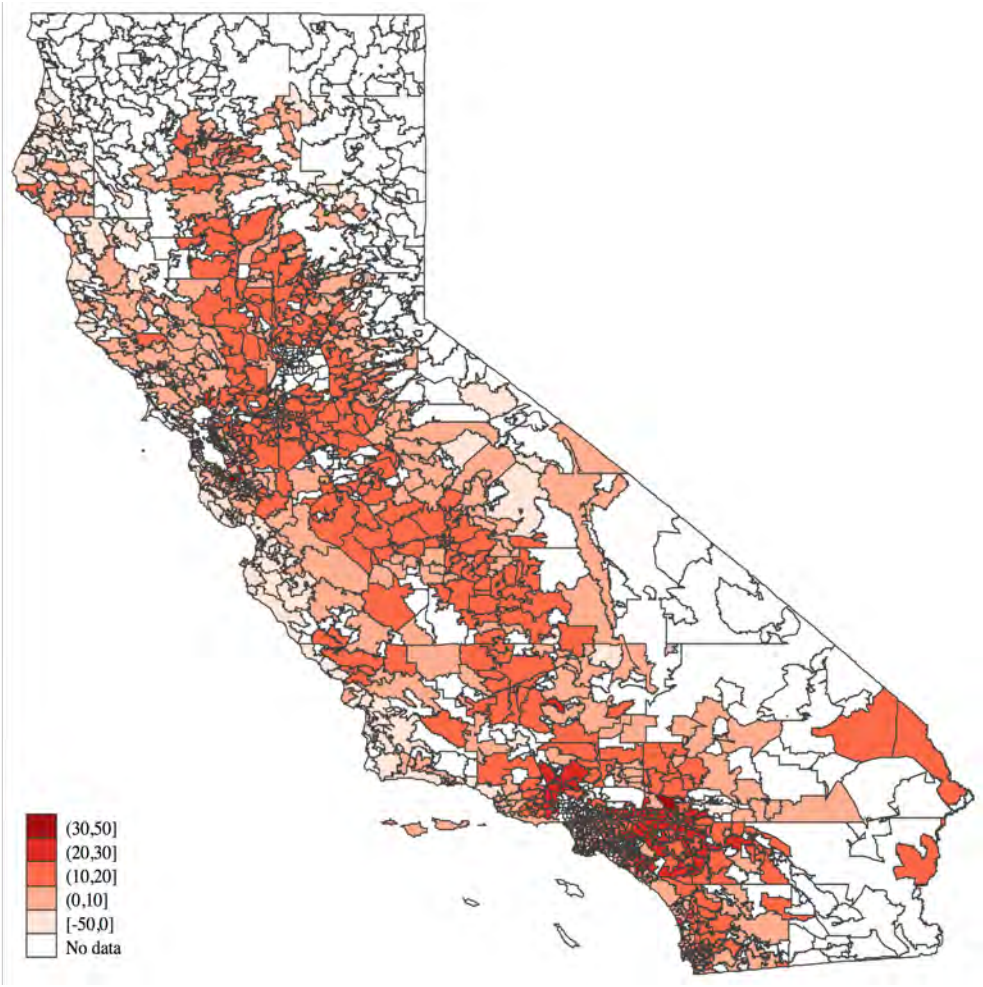
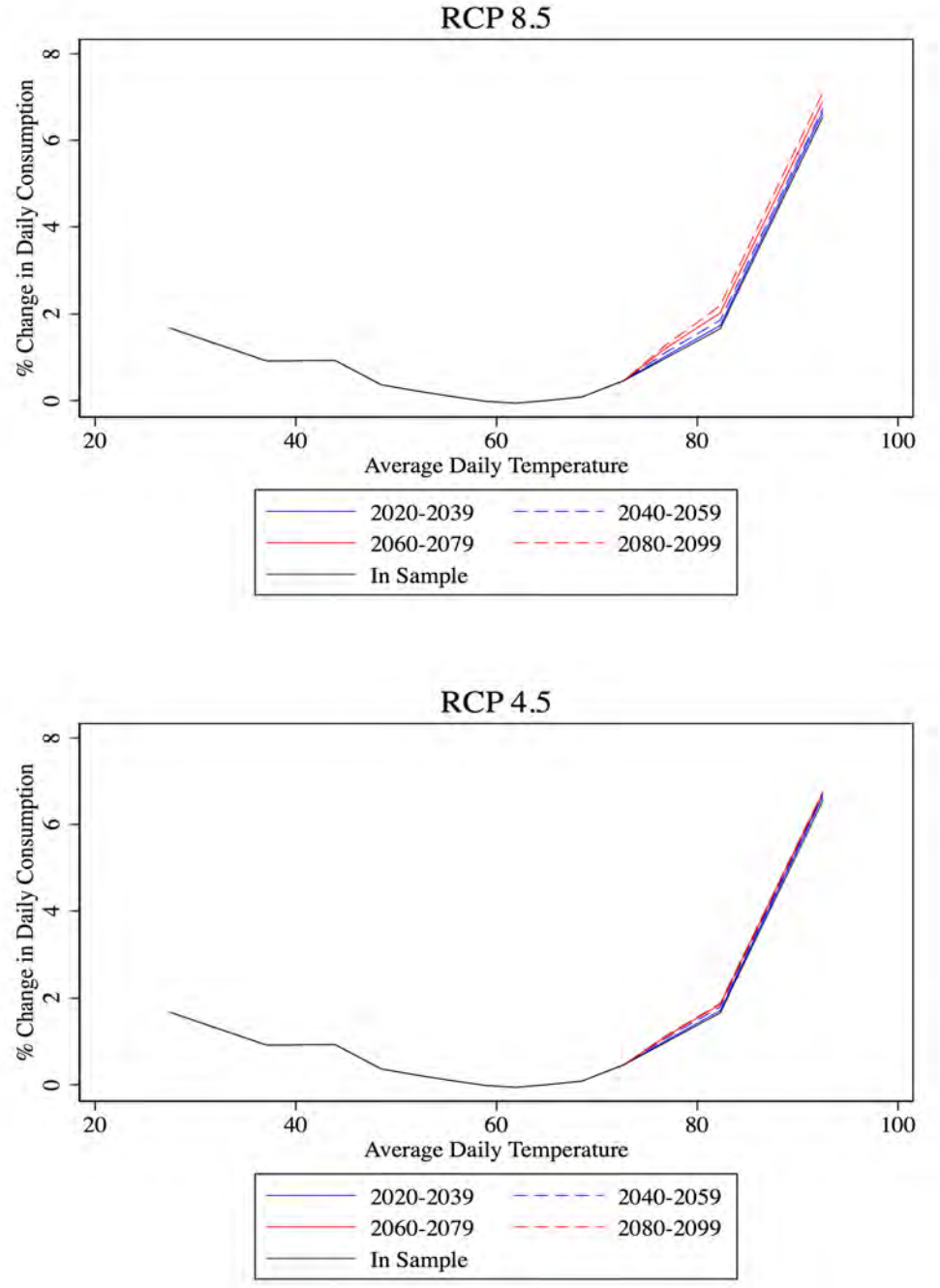


Figure 5: Intensive Margin Adjustment: Projected Percent Increases in Average Household Electricity Consumption 2080-2099 over 2000-2015



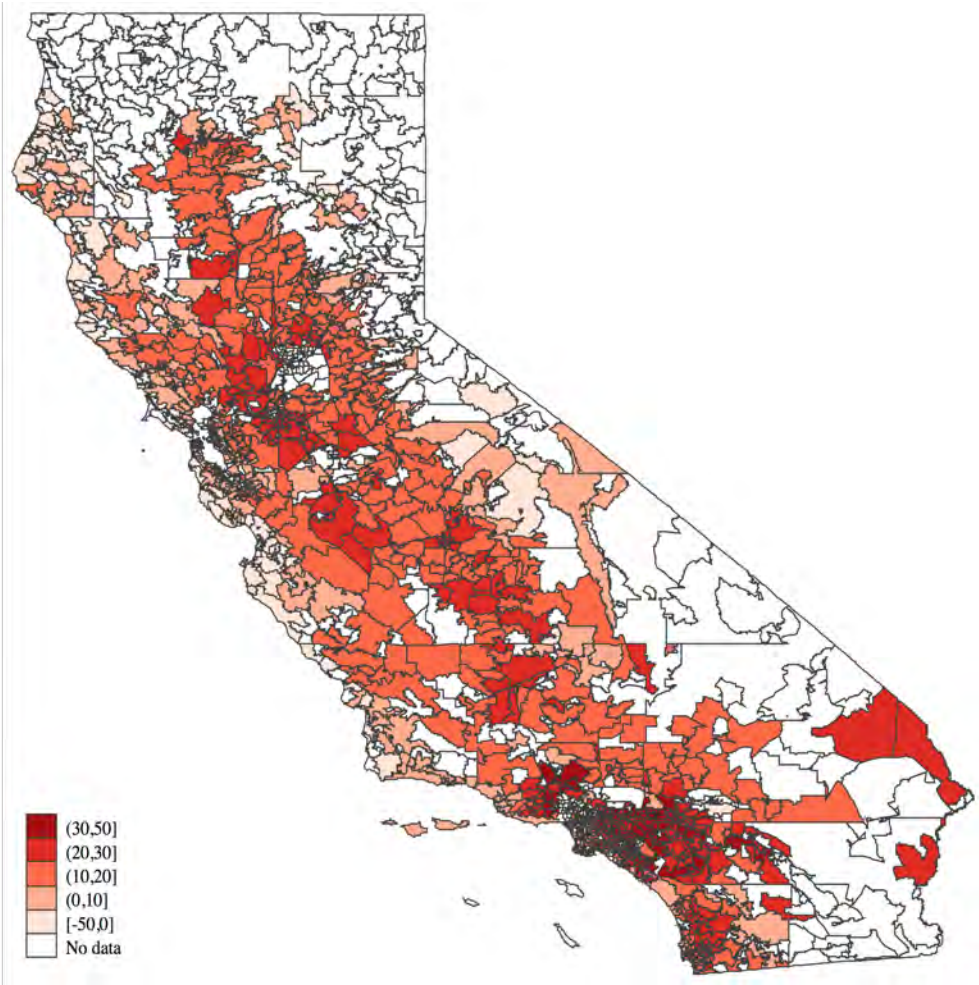
Notes: This figure plots the average per household increase across all 18 GCMs for RCP8.5 for the last two decades of this century over the years 2000-2015. The figure holds the temperature response curve fixed at the values estimated in-sample.

Figure 6: In Sample and Simulated Future Population-Weighted Temperature Response Curves.



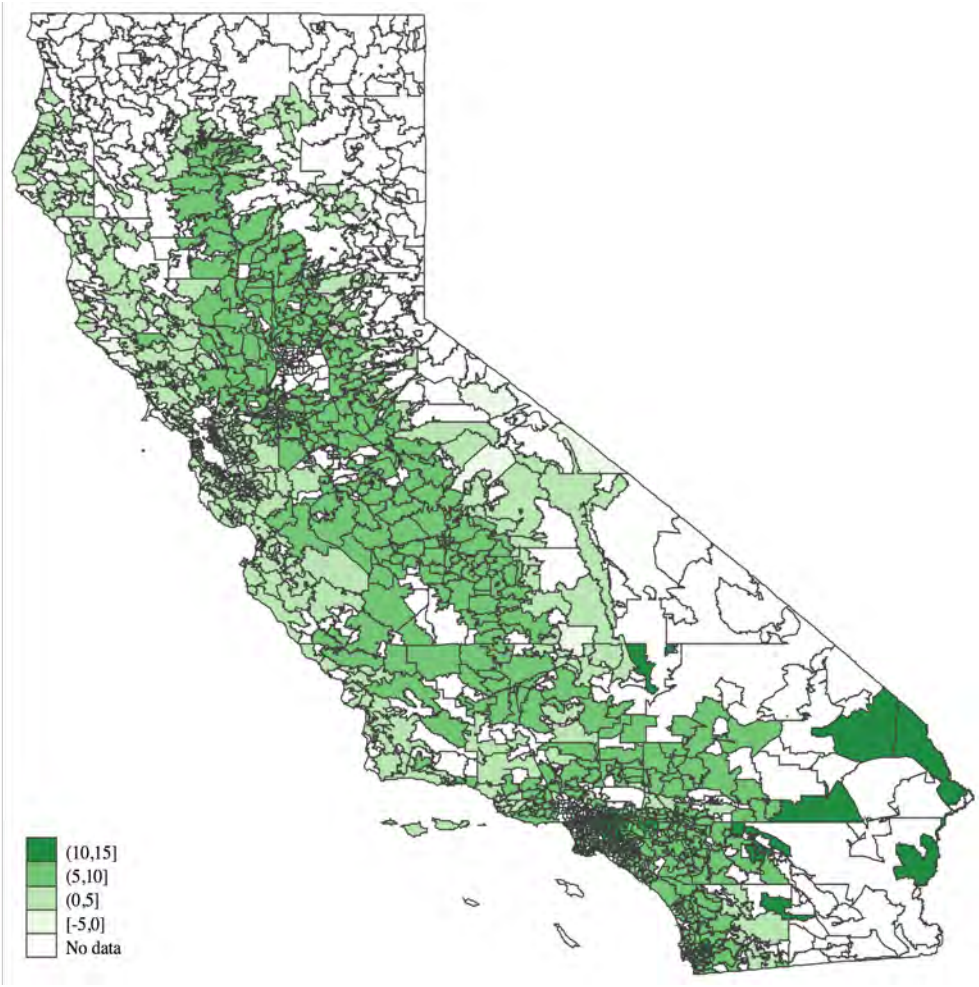
Notes: This figure plots the population weighted average of the temperature response curves across all 18 GCMs (climate models) in blue and red. The solid black line displays the in-sample estimated population weighted average across all zip codes temperature response curve.

Figure 7: Extensive Margin Adjustment: Projected Percent Increases in Average Household Electricity Consumption 2080-2099 over 2000-2015



Notes: This figure plots the average per household increase across all 18 GCMs for RCP8.5 for the last two decades of this century over the years 2000-2015. The figure varies both the slope of the response curve and movement along the new ZIP level response curves allowing for both extensive and intensive margin adjustment.

Figure 8: Difference between Extensive nd Intensive Margin Impacts on Average Household Electricity Consumption 2080-2099 over 2000-2015



Notes: This figure plots the difference between figures 6 and figure 5 to display the additional impact from the extensive margin adjustments.

Table 2: Summary Statistics

		Usage Per Bill		Average Price		Percentage of Days in Percentiles of			
CEC		(kWh)		($/kWh)		% Days with Average Temperature			
Climate Zone	Bills	Mean	s.d.	Mean	s.d.	< 32°F	[32, 42) °F	[79, 86) °F	> 86°F
1: Arcata	2,192,756	572	339	0.13	0.03	0.0%	3.6%	0.0%	0.0%
2: Santa Rosa	17,300,000	616	359	0.13	0.02	0.1%	4.0%	1.2%	0.1%
3: Oakland	75,600,000	479	291	0.13	0.02	0.0%	0.4%	0.2%	0.0%
4: Sunnyvale	32,600,000	603	337	0.13	0.02	0.0%	1.2%	1.5%	0.1%
5: Santa Maria	7,754,079	506	304	0.13	0.03	0.0%	0.2%	0.1%	0.0%
6: Los Angeles	73,900,000	530	315	0.13	0.03	0.0%	0.0%	0.6%	0.0%
7: San Diego	76,200,000	503	303	0.15	0.03	0.0%	0.0%	0.7%	0.0%
8: El Toro	81,800,000	577	331	0.14	0.03	0.0%	0.0%	2.0%	0.1%
9: Pasadena	72,000,000	614	353	0.13	0.02	0.0%	0.3%	5.3%	0.3%
10: Riverside	86,000,000	674	374	0.14	0.03	0.1%	0.4%	8.5%	0.6%
11: Red Bluff	12,700,000	761	400	0.13	0.02	0.1%	4.3%	8.1%	1.3%
12: Sacramento	50,600,000	696	374	0.13	0.02	0.1%	2.9%	5.6%	0.7%
13: Fresno	33,100,000	727	409	0.13	0.02	0.1%	3.1%	14.0%	3.3%
14: China Lake	17,900,000	687	375	0.13	0.03	0.1%	5.2%	11.2%	4.4%
15: El Centro	8,534,004	674	465	0.13	0.03	0.0%	1.1%	12.8%	15.5%
16: Mount Shasta	10,100,000	602	369	0.13	0.02	5.2%	17.2%	2.5%	0.1%

Note: The table displays summary statistics for residential electricity consumption for the sample used in the estimation. s.d. = standard deviation.

Table 3: Summary Statistics for ZIP Codes In and Out of Sample

Variable	In Sample			Not In Sample			
	n	Mean	Std. Dev.	Obs	Mean	Std. Dev.	t p-value
Population ('000)	1235	25.08	21.17	517	12.15	18.82	<0.001
Share White	1235	0.70	0.20	517	0.75	0.19	<0.001
Share Afr. Am.	1235	0.05	0.08	517	0.05	0.07	0.071
Share Hispanic	1235	0.30	0.23	517	0.25	0.24	<0.001
Share Asian	1235	0.11	0.14	517	0.08	0.10	<0.001
Share Hawaiian	1235	0.01	0.01	517	0.01	0.01	0.668
Share Nat. American	1235	0.03	0.03	517	0.04	0.07	<0.001
Share Male	1235	0.50	0.03	517	0.52	0.09	<0.001
Person per Household	1235	2.82	0.61	517	2.56	0.98	<0.001
Home Value (10kUS$)	1235	39.85	23.99	517	28.21	25.32	<0.001
HH Income (1kUS$)	1235	64.83	27.41	517	46.73	33.38	<0.001
Median Age	1235	39.01	8.01	517	40.90	10.02	<0.001
Elevation	1235	371.27	607.70	517	897.03	1329.93	<0.001
JJA Tavg 1950-2010	1235	71.52	5.25	517	71.25	7.37	0.385

Table 4: Second Stage Regressions of Temperature Response Coefficients by Temperature Bin

	(1)	(2)	(3)	(4)
Historical Bin Tavg Share	-0.053***	-0.0526***	0.0051***	0.0012***
	(0.000)	(0.000)	(0.000)	(0.000)
Interaction Bin 12+				0.0628***
				(0.000)
Income	No	Yes	Yes	Yes
Bin Fixed Effects	No	No	Yes	Yes
Observations	5,612	5,267	5,267	5,267

Table 5: Projected Percent Increase in Residential Electricity Consumption due to Climate Change over 2000-2015 Consumption Averaged Across Climate Models by Forcing Scenario, and Extensive Margin Adjustment Modeling

Simulation	(1)	(2)	(3)	(4)
Households	Normal	Normal	Normal	Normal
RCP	4.5	8.5	4.5	8.5
Intensive Margin	Yes	Yes	Yes	Yes
Extensive Margin	No	No	Yes	Yes
2020 - 39	1.19%	1.45%	0.99%	1.37%
2040 - 59	2.44%	3.40%	2.86%	4.37%
2060 - 79	3.37%	6.78%	4.33%	10.01%
2080 - 99	3.85%	10.82%	5.12%	17.27%

Table 6: Joint Impacts of climate change and population growth

Population Growth Scenario	Low	Low	Medium	Medium	High	High
RCP	4.5	8.5	4.5	8.5	4.5	8.5
Extensive Margin	Yes	Yes	Yes	Yes	Yes	Yes
2020-39	15%	15%	39%	40%	55%	55%
2040-59	18%	19%	66%	68%	103%	105%
2060-79	21%	27%	93%	99%	171%	177%
2080-99	28%	40%	123%	135%	277%	289%

Note: This table displays the simulated percent increase in total residential electricity consumption relative to 2010 population and 2006-2015 climate for low, medium and high population growth scenarios using both intensive and extensive margin adjustments. The figures are population weighted averages across climate models. The population weights are held constant at 2010 levels.

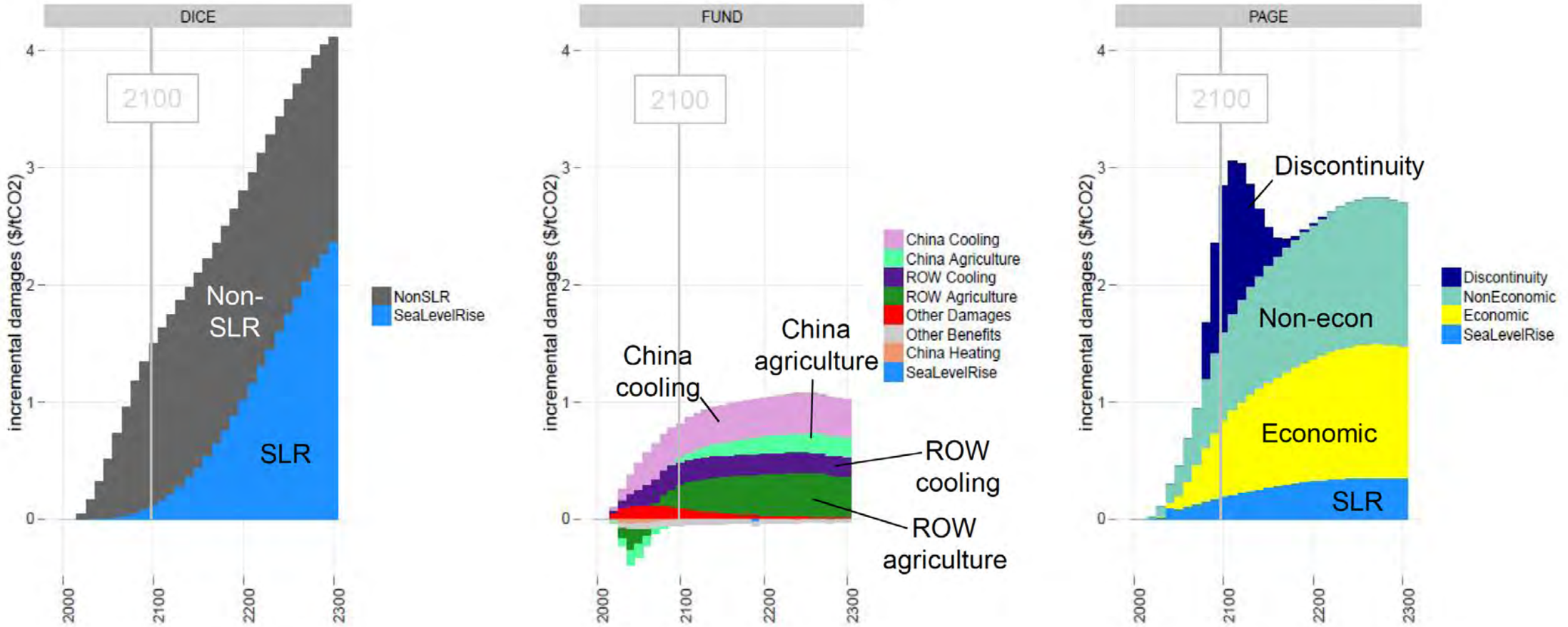
Quantifying the Long Run Impacts of Climate Change on Electricity and Natural Gas Consumption Using Big Data

Maximilian Auffhammer
University of California at Berkeley
National Bureau of Economic Research & CESifo

The best camp evah, July 1, 2016

Berkeley
UNIVERSITY OF CALIFORNIA

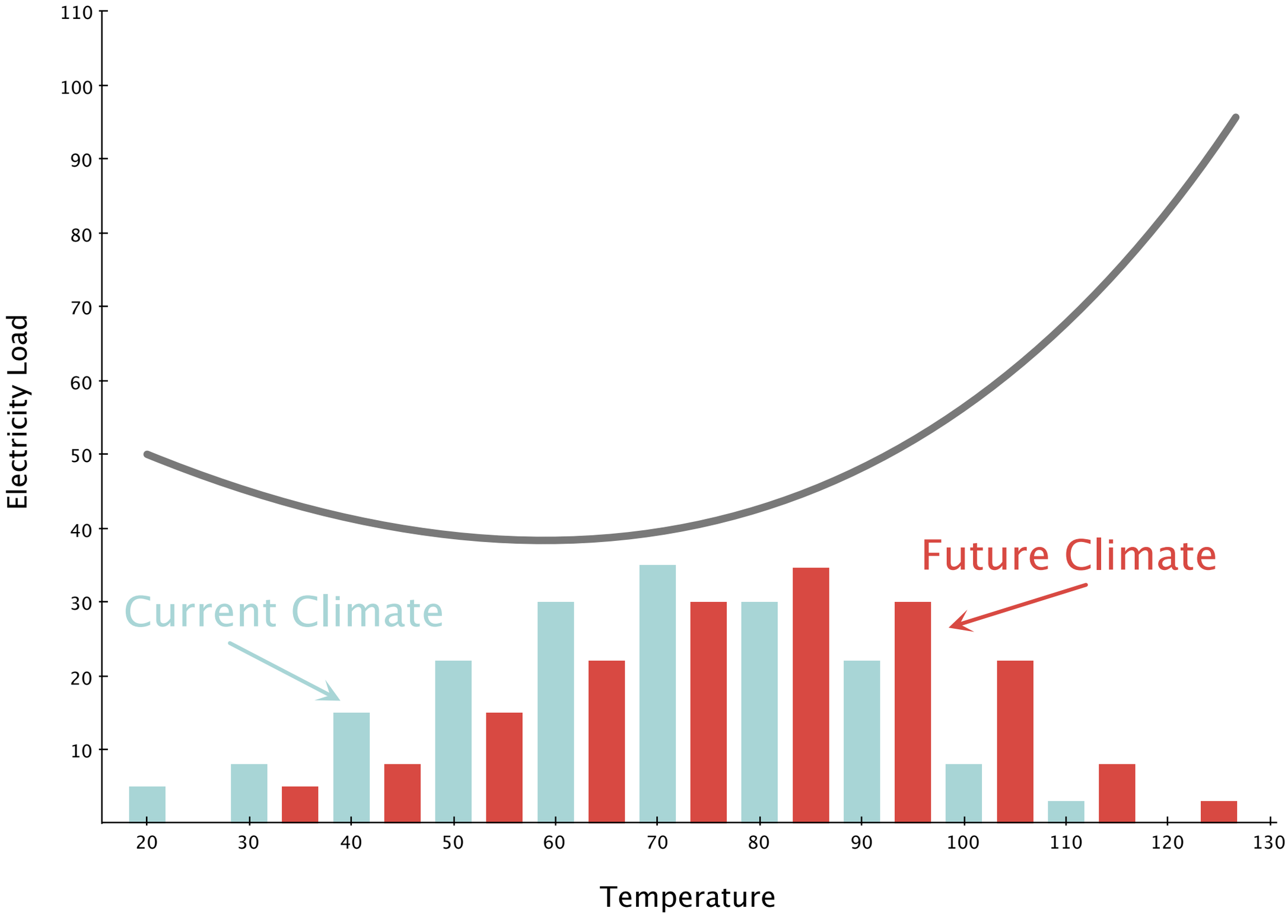
Rose et al. (2015) comparison of models



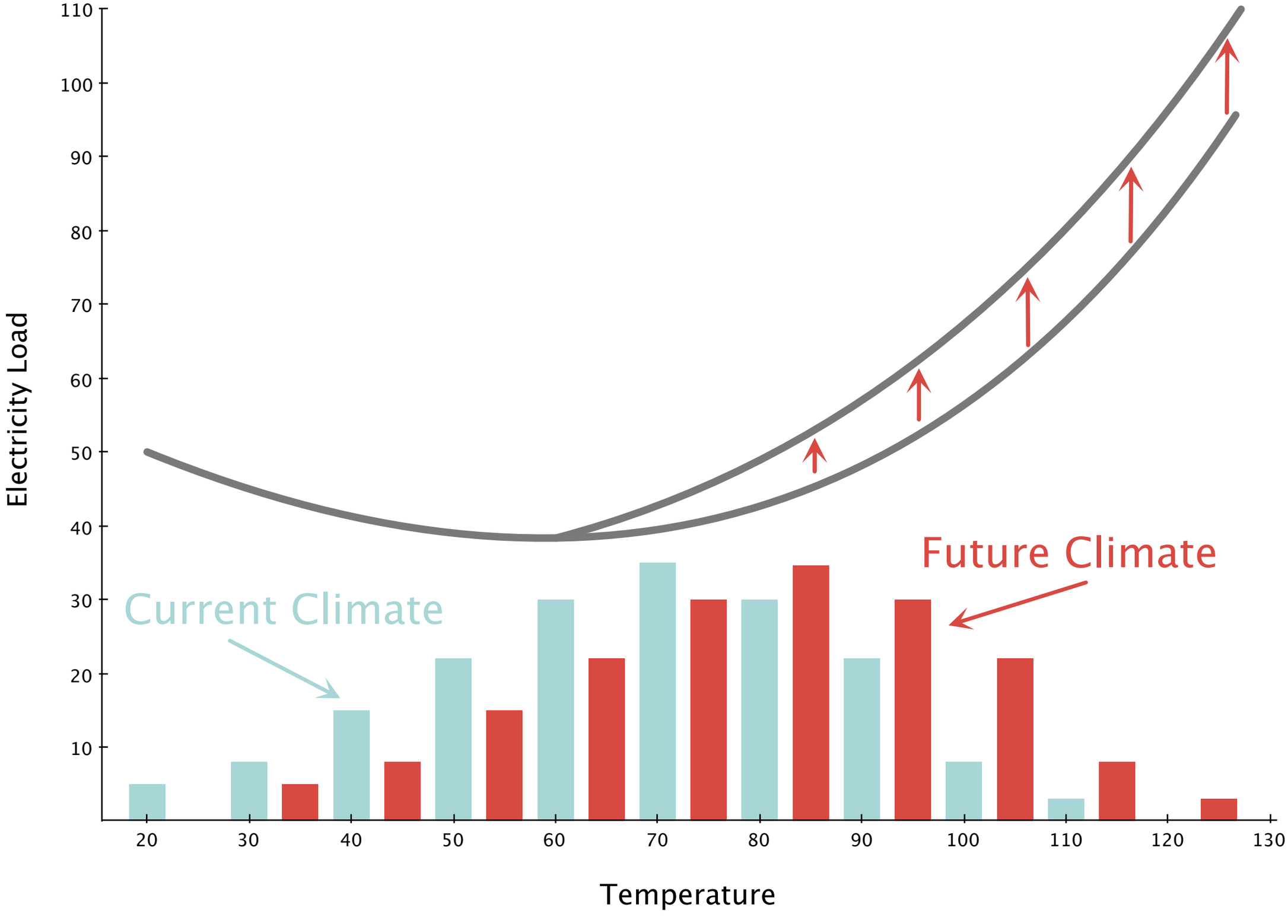
- In the FUND model the space cooling impacts are calculated from this equation:

$$SC_{t,r} = \alpha_r Y_{1990,r} \left(\frac{T_t}{1.0}\right)^{\beta} \left(\frac{y_{t,r}}{y_{1990,r}}\right)^{\epsilon} \left(\frac{P_{t,r}}{P_{1990,r}}\right) / \prod_{s=1990}^{t} AEEI_{s,r}$$

Operating existing air conditioners more increases load



Changing climate moves the temperature load curve



Why looking at the long run response is interesting.

Model 1.0: Univariate Time Series Models - Short run climate elasticity (e.g. Franco and Sanstad 2008)

- Exploit time series variation in weather
- Sensitive to any univariate time series issue there ever was

Model 2.0: Ricardian Models - Long run climate elasticity (e.g. Mendelssohn et al. 19xx/20xx)

- Exploit cross section in climate
- Sensitive to omitted variables bias

Model 3.0: Panel Models - Short run weather elasticity (e.g. Auffhammer, Ramanathan and Vincent (2006), DG (2007))

- Exploit random nature of weather shocks
- Estimate a weather sensitivity

Model 4.0: Long Difference Models - Short and long run elasticity (e.g. Burke and Emerick, 2015)

- Exploit observed warming variation across counties
- Require long panels outcomes and temperatures

Research Question

How much will residential electricity consumption change by the end of the current century in a world with climate change

- holding air conditioner penetration constant (intensive margin)
- letting air conditioner penetration change (extensive margin)

Preview of approach and findings

Model 5.0: Climate **A**daptive **R**esponse

- Use a jealousy inducing dataset on household electricity and natural gas consumption to estimate causal short run response to temperature at ZIP code level.
- Explain cross sectional variation in electricity temperature response across ZIP codes as a function of climate.
- Link to state of the art downscaled climate model and calculate projections of impacts of short and long run response

Main findings:

- Accounting for extensive margin adjustments roughly triples moderate impacts
- Spatial pattern of intensive margin and extensive margin adjustment is heterogenous.

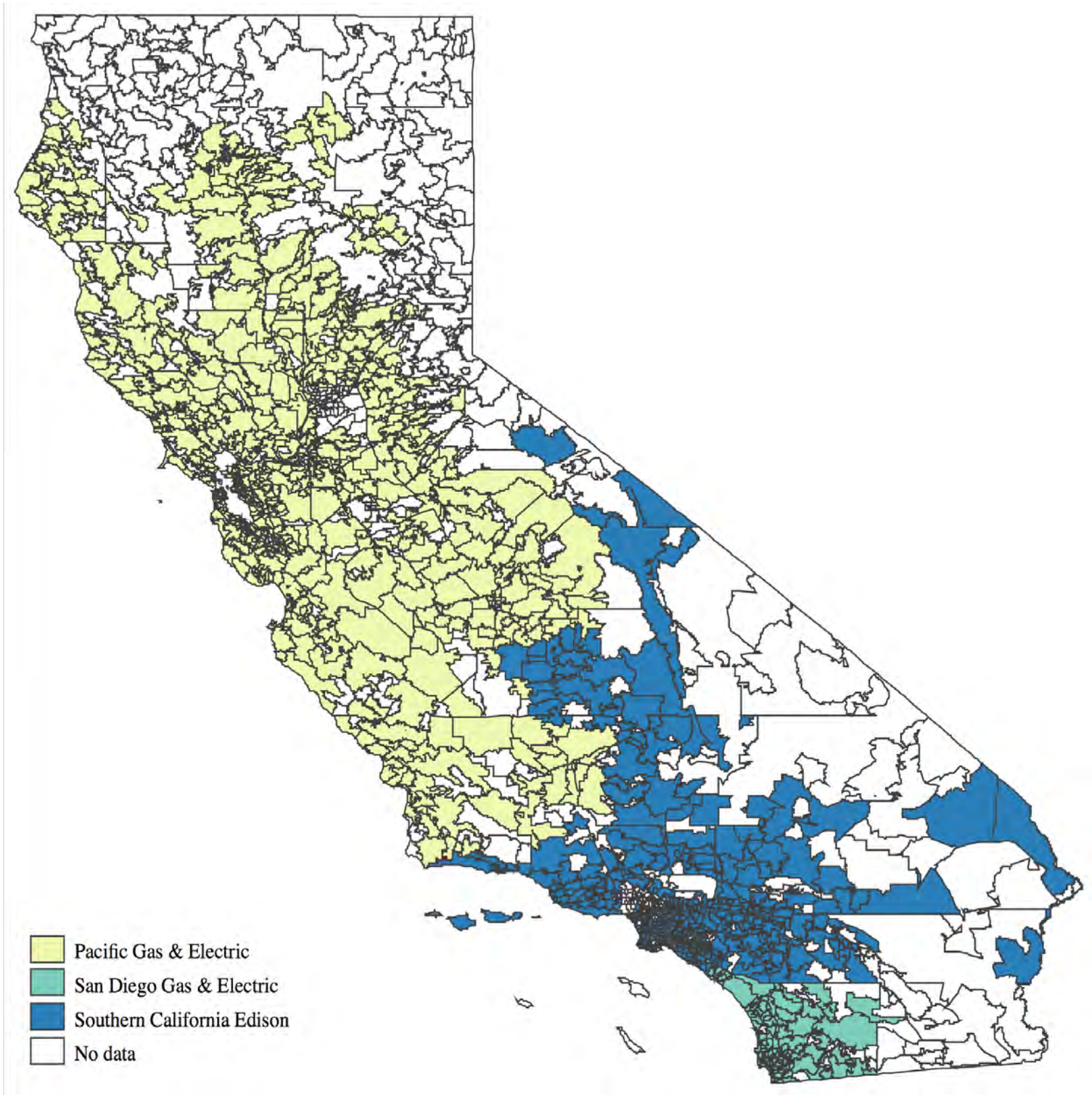
The billing data

- Complete residential gas and electric billing data for California's investor owned utilities (PG&E, SCE, SDG&E)
- ~82% of all California households
- Drop bills with daily consumption less than 2 kWh
- Drop bills for solar households
- Electricity Dataset has 914 million bills:
  - **658 million "normal" bills**
  - 164 million "CARE" bills
  - 92 million "all-electric" bills

The billing data

Utility	*Energy Type*	*Coverage*	*~ Number of Bills*
PG&E	Electricity	2003-2009	342 Million
PG&E	Natural Gas	2004-2014	587 Million
SCE	Electricity	1999-2008	469 Million
SoCalGas	Natural Gas	2010-2015	267 Million
SDG&E	Electricity	2000-2009	153 Million
SDG&E	Natural Gas	2008-2015	74 Million

The billing territory covers most of the state



Estimation Strategy

First Stage: Intensive Margin Response estimated separately for each ZIP Code *j*:

$$\log(q_{it}) = \sum_{p=1}^{14} \beta_j^p D_{pit} + \gamma Z_{it} + \alpha_i + \phi_m + \psi_y + \varepsilon_{it}$$

Second Stage: Extensive Margin Response estimated across ZIP Codes

$$\beta_j^p = \delta_1 + \delta_2 C_j^p + \boldsymbol{\delta_3 Z_j} + \eta_j$$

Functional forms in first and second stage

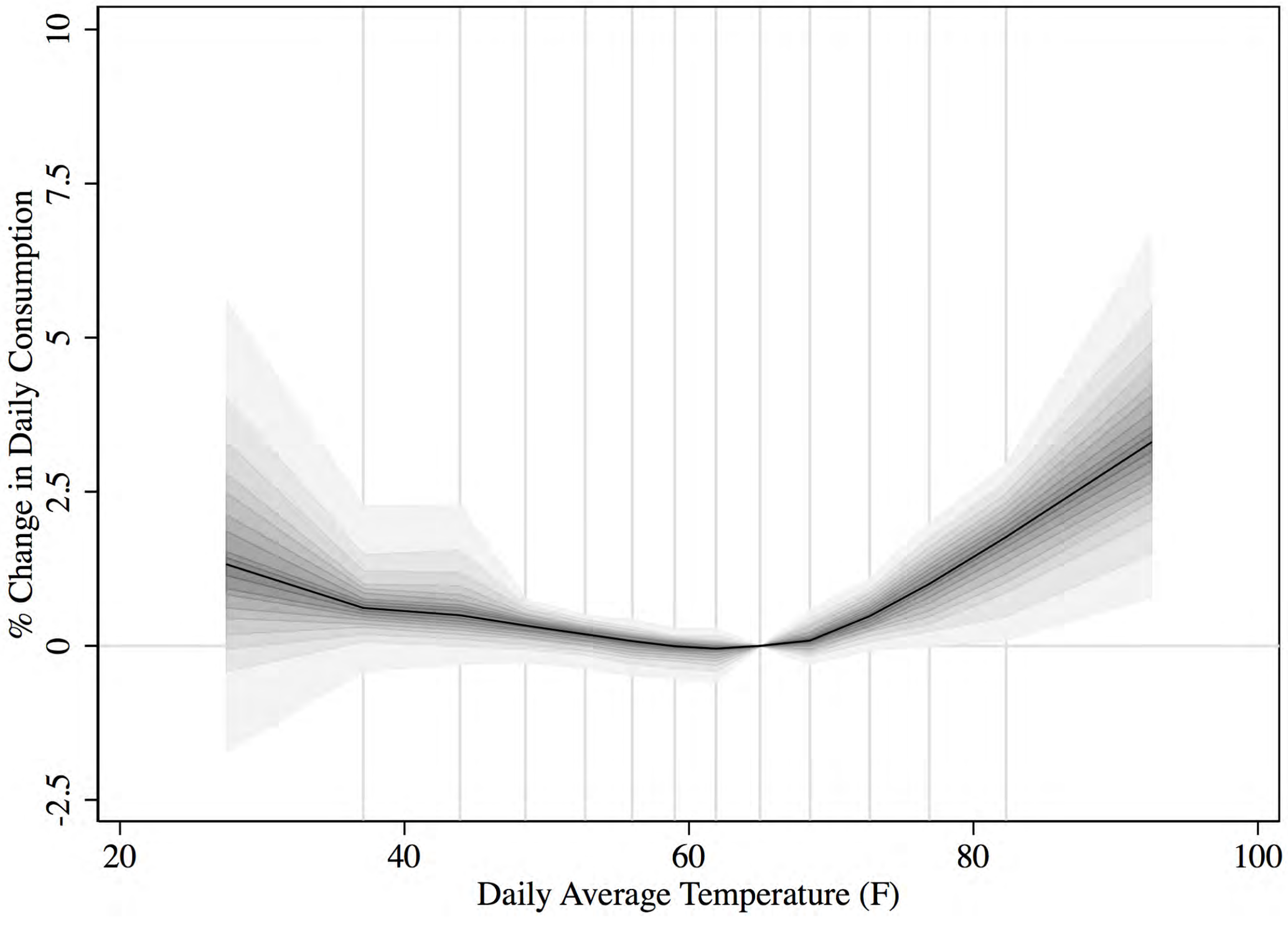
Flexible functional form for temperature in **first stage** :

- D_{pit}= number of days of an electricity bill in p^{th} temperature bin
- Percentile bins approach:
- 14 bins with cutoff at 1^{st} , 5^{th} ,10^{th} ,....,90^{th} ,95^{th},99^{th} percentile
- These translate into cutoffs of 32, 42, 46, 51, 54.5, 58, 60, 63, 67, 70, 75, 79 and 86 degrees Fahrenheit

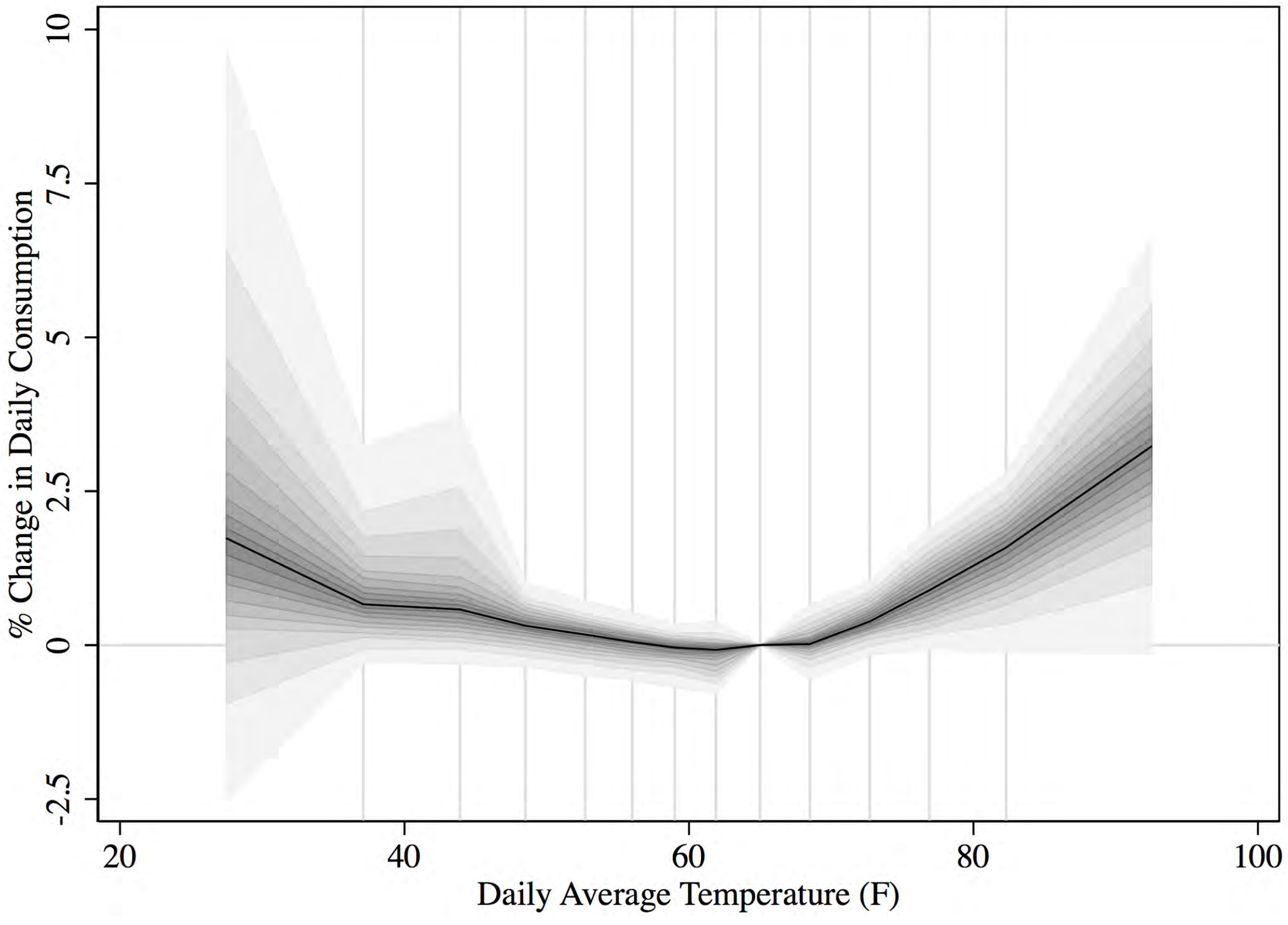
Climate impact in **second stage** :

- C_j^p= Share of days spent in bin p during 2000-2010

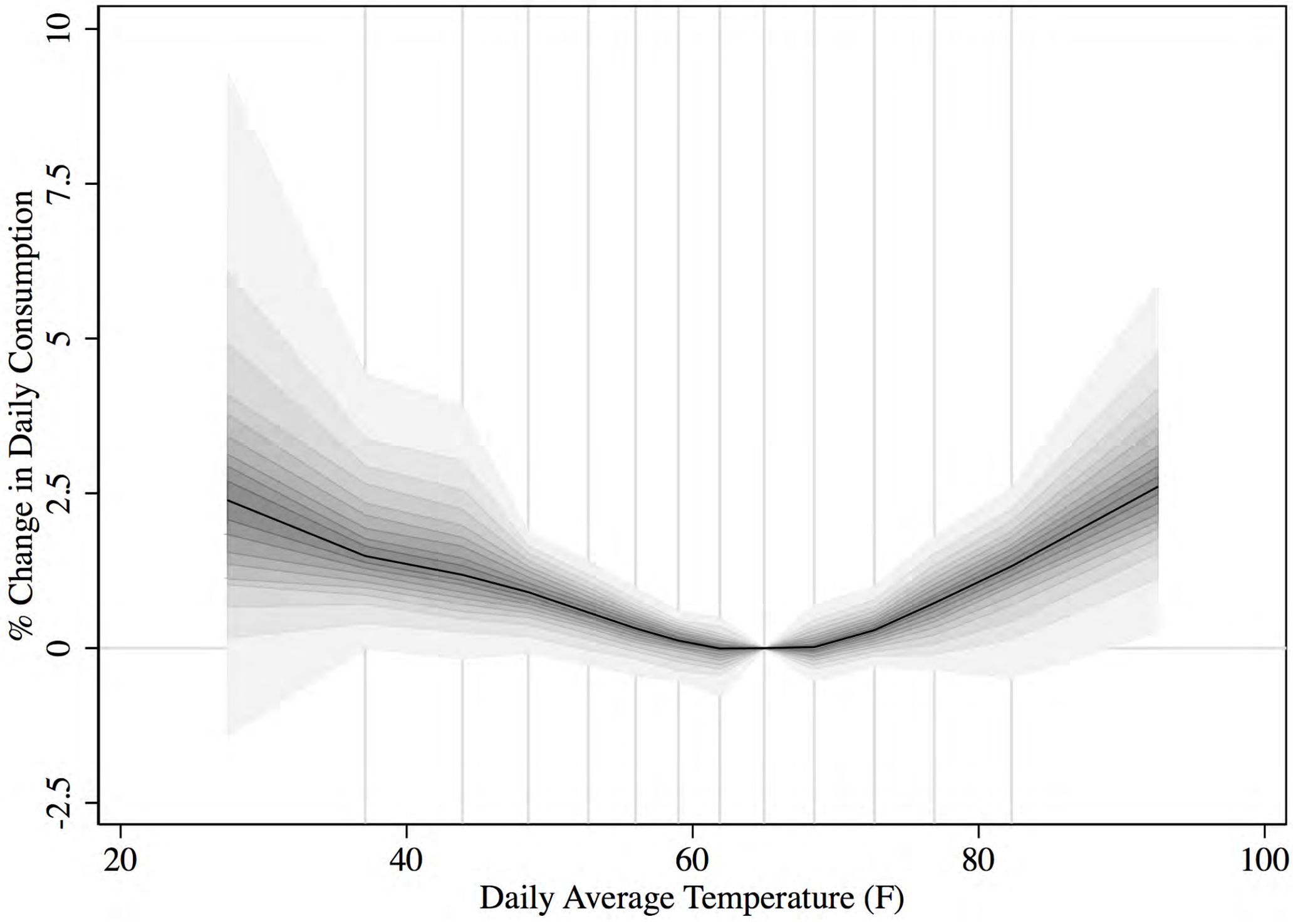
Temperature Response: “Normal households”



All “CARE” households



All “Electric” households

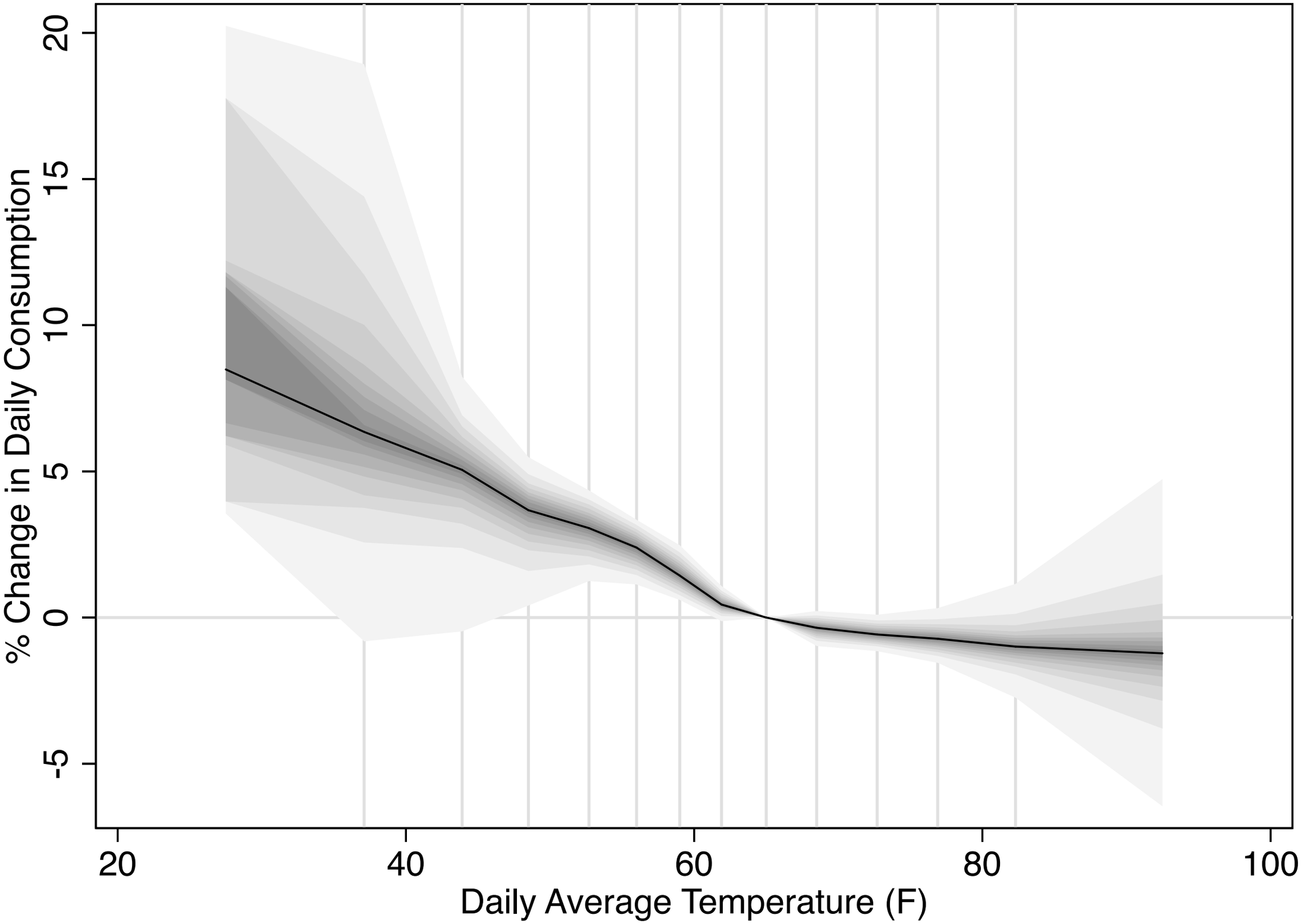


Second stage regression

Table 3: Second Stage Regressions of Temperature Response Coefficients by Temperature Bin

	(1)	(2)	(3)	(4)
Historical Bin Tavg Share	-0.053***	-0.0526***	0.0051***	0.0012***
	(0.000)	(0.000)	(0.000)	(0.000)
Interaction Bin 12+				0.0628***
				(0.000)
Income	No	Yes	Yes	Yes
Bin Fixed Effects	No	No	Yes	Yes
Observations	5,612	5,267	5,267	5,267

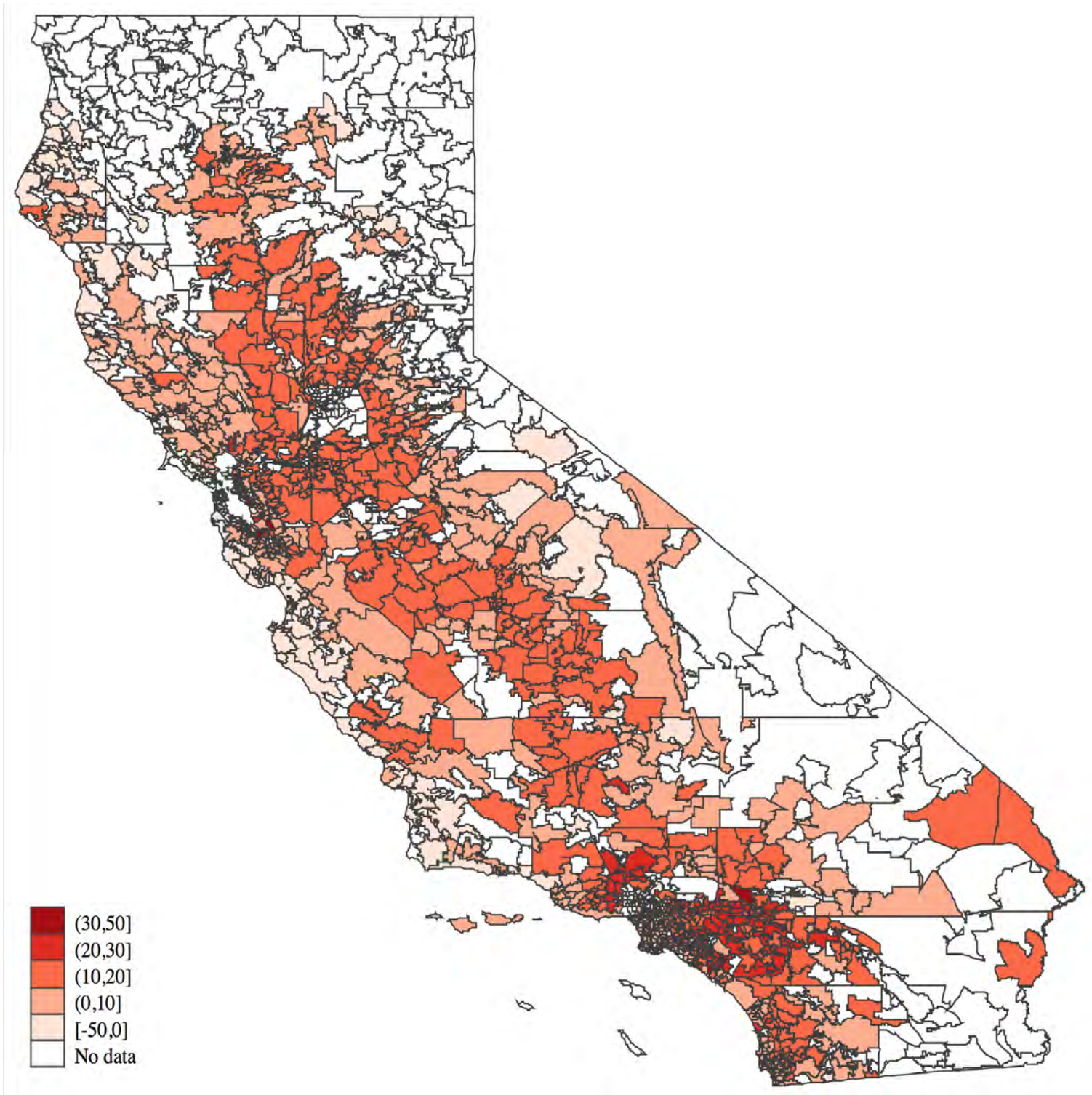
Natural Gas Response



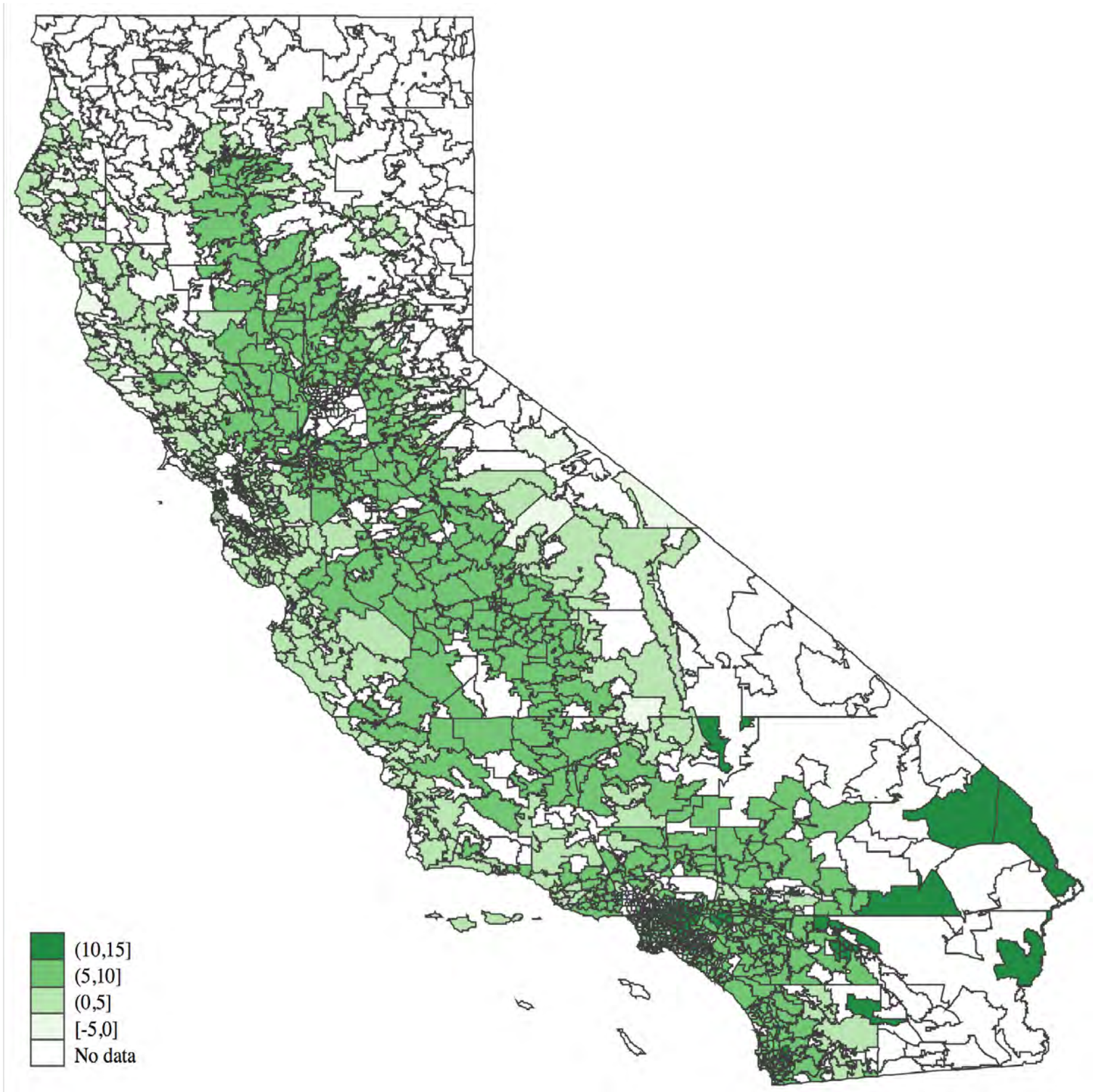
Simulated Impacts: GCM

- Output from 18 climate models from the IPCC's 5th assessment report
- Each model provided for RCP 4.5 and RCP 8.5
- Model output processed:
  - Downscaled
  - Daily
  - Rainfall, TMin, TMax
- Climate data are processed using the same algorithm used to process weather data.
- Doing this is really painful. To an economist.

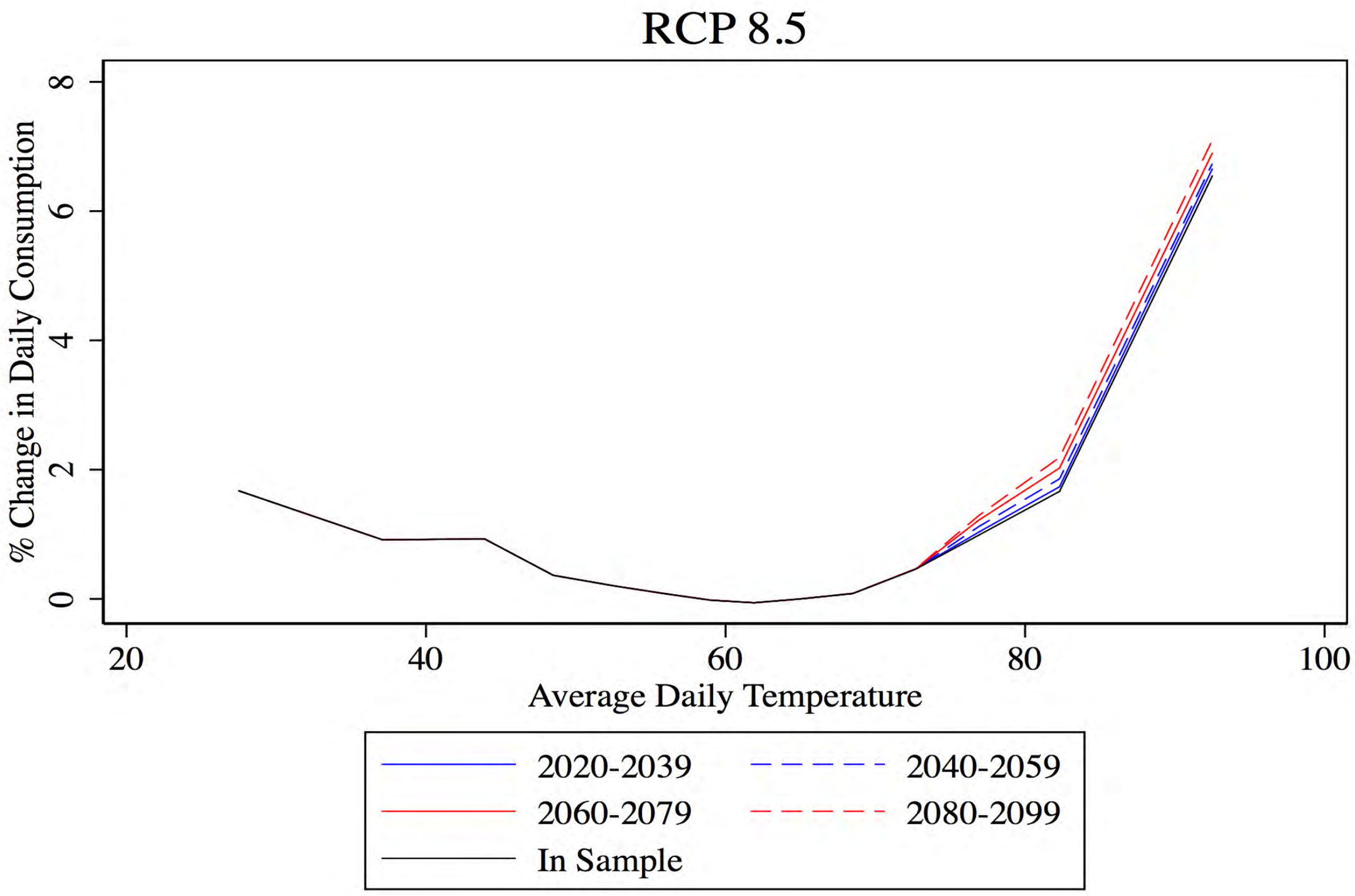
Projected intensive margin impacts



Projected extensive margin impacts



Population weighted state-wide response function



Projected aggregate temperature impacts in %

Simulation	(1)	(2)	(3)	(4)	(5)	(6)
Fuel	Elec	Elec	Elec	Elec	Gas	Gas
Extensive?	No	No	Yes	Yes	No	No
RCP	4.5	8.5	4.5	8.5	4.5	8.5
2020-2039	1.2	1.5	1.0	1.4	-3.7	-4.6
2040-2059	2.4	3.4	2.9	4.4	-7.5	-9.8
2060-2079	3.4	6.8	4.3	10.0	-9.7	-15.4
2080-2099	3.9	10.9	5.1	17.3	-10.7	-19.6

Conclusions

- I know I am so out of time by now. But I need to give you a take home message. So here it is.
- California Homes used 0.287 quadrillion BTU of electricity and 0.439 quadrillion BTU of natural gas in 2009 (EIA, RECS).
- You see where I am going with this? Yeah you are.
- Climate Change is simulated to lead to a **0.036 quad BTU net decrease** in energy consumption for the residential sector in California
- Of course, everywhere else is different from California
- Consumption changes are not the only costs.

But, we (EI@Haas) do not only produce pessimistic results.

