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

February 26, 2017

Abstract

This paper proposes a simple method (**C**limate **A**daptive **R**esponse **E**stimation - CARE) to parameterize the climate dose response function using residential electricity and natural gas demand for the world's seventh largest economy - California. The advantage of the proposed method is that it only requires detailed information of consumption, yet does not require knowledge of what technology is installed. Using almost two billion energy bills, we estimate spatially highly disaggregated intensive margin temperature response functions (e.g. increases in air conditioning) using daily variation in weather. In a second stage we explain variation in the slope of the dose response functions (e.g. the adoption of additional air conditioners) across space as a function of summer climate. Using 18 state of the art climate models we simulate future demand by letting households vary consumption along the intensive and extensive margin. 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 the future impacts on electricity consumption.

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. This paper is based on the 2012 PIER Report "Hotspots of climate-driven increases in residential electricity demand: A simulation exercise based on household level billing data for California". Joe Shapiro and seminar participants at Stanford University, UC Irvine, Iowa State University, UCSB, UIUC, Lawrence Berkeley Labs, Arizona State University, Resources for the Future and the 2017 ASSA Meetings have provided valuable feedback. All remaining errors are the authors.

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°C (1.8 - 6.7°F) 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). Humans respond to hot outdoor ambient temperatures by cooling the indoor environment at home and/or at work. If the frequency and intensity of hot days increases due to climate change, one would expect this to cause increased cooling and decreased heating demand. One of the three Integrated Assessment Models used in the calculation of the federal Social Cost of Carbon concludes that increased space cooling is the major driver of Global Climate Damages (Rose et al, 2014). This finding relies on an assumed temperature responsiveness of a simple space cooling function in the FUND Integrated Assessment model, which has little to no empirical basis.

Air conditioning is the main adaptation mechanism open to humans and has been shown to be an 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 due to increased penetration of air conditioners. The observed trajectory of air conditioner installation has been driven by growth in incomes and falling prices of AC units and the electricity required to operate them over the past century (Biddle, 2008). Hence areas, which have a hotter climate, are wealthier and have lower electricity costs have higher levels of air conditioner penetration. (EIA, 2011). However, a changing climate represents a new driver of air conditioner adoption. If San Francisco with its pleasant coastal climate by end of century "receives" Fresno's hot climate, even San Franciscans will install window units in existing apartments and new construction will have central air conditioning units installed. The cost of this adaptation mechanisms comes in the form of installation and operating costs, while the benefits accrue in the form of better health outcomes and comfort. A recent review of the literature (Auffhammer and Mansur, 2014) lamented the dearth of causal estimates of empirically calibrated damage functions, which quantify the short and long run 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 climate driven 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 the United States. Davis and Gertler (2015) is the only example for a large country which utilizes data both on appliance holdings and electricity consumption for a large share of the population.

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 consumption when one *does not have data on installed capital* (e.g. air conditioners). In a first stage, we estimate the causal 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 more than 1,000 ZIP codes in California in our sample. Estimation at this fine level is made possible by the fact that we observe almost 2 billion electricity and natural gas bills which represent 79 percent of California's household over a decade.

For electricity, 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 average weather ("climate"), income and population density. Technically speaking, in this "second stage" we regress the slope of each ZIP code's temperature response function in different temperature bins on income, population density and summer climate. We therefore separate the impact of income and population density 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 climate projections 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 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 quantifying the economic impacts of climate change has exploded over the past decade. Review articles by Carleton and Hsiang (2016), Hsiang (2016), and Dell et al. (2014) provide up to date and comprehensive overviews of both methods and applications. The key challenge is to estimate externally valid dose response functions between economic outcomes of interest (e.g. energy consumption, crop yields, mortality, water consumption, labor productivity, cognitive ability) and a long (e.g. 30 year) average of weather, which is commonly referred to as climate. This estimated long run response is supposed to capture both adaptation behavior at the intensive margin (e.g. increased operation of existing air conditioners) and extensive margin (e.g. installation of additional air conditioners). The coefficients parameterizing said dose response function should be estimated in a way that allows them to take on a causal interpretation. This is anything but straightforward. Below I provide a brief summary of the methodological approaches in existing papers, while listing examples with an energy focus. For a broader review, consult Carleton and Hsiang (2016).

The earliest literature relied on large-scale bottom-up simulation models to simulate future electricity demand under varying climate scenarios. The advantage of these models is that they can 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 early papers in this literature suggest that climate change will significantly increase energy consumption (Cline, 1992; Linder et al. (1987), Baxter and Calandri (1992); Rosenthal et al. (1995)).

The recent literature has focused on providing empirical estimates of climate response functions for a large number of sectors. There are four empirical approaches using distinctly different sources of variation to parameterize climate response functions: (1) Time Series Regression (2) Ricardian Approach (3) Panel Estimation (4) Long differences. Each of these approaches has distinct advantages and disadvantages. This paper builds on insights in other papers to provide a fifth approach, which I call Climate Adaptive Response Estimation (CARE).

A simple and commonly practiced approach taken to quantify the impact of climate on electricity consumption uses high frequency (e.g. daily or hourly) time series of electricity load and regresses these on population weighted functions of weather. Franco and Sanstad (2008) use hourly electricity load for the entire California ISO over the course of the year 2004 and regress them on average daily temperature. They identify a highly nonlinear response of load to temperature. They show projected increases in electricity consumption and peak load of 0.9 to 20.3 percent and

1.0 to 19.3 percent, respectively. Crowley and Joutz (2003) use a similar approach for the PJM Interconnection. Auffhammer et. al (2017) estimate the response of peak load and average load to daily weather for 166 load balancing authorities, covering the vast majority of US electricity load. They show modest increases in consumption by the end of this century, yet significant increase in the intensity of annual peak load (15-21%) and a twelve to fifteen fold increase in peak events by the end of century. The drawback of this approach is that it relies on short term fluctuations in weather and hence does not estimate a long run climate response but rather a short run weather response. It simply cannot account for adaptation responses to climate change such as increased use and installation of air conditioners or increased incidence of demand side management and energy efficiency programs.

The second strand of the literature is based upon the seminal work by Mendelsohn et al. (1994), who estimated the impact of climate change on agricultural yields by regressing yields or net profits on climate. This cross sectional approach has the advantage that it estimates a true climate response. The method has been widely criticized, as any non-experimental cross sectional regression is bound to suffer from omitted variables bias (e.g. Deschenes and Greenstone, 2007). Any unobserved factor correlated with climate and the outcome of interest will bias the coefficients on the climate variable. This approach has not been widely applied in the energy literature, yet Mansur et al. (2008) is one example of cross sectional approaches are Mansur et al. (2008) and Mendelsohn (2003). The innovation in these papers is that they endogenize fuel choice, which is usually assumed to be exogenous and provides one avenue of adaptation.

The third strand of the literature relies on panel data of energy consumption at the household, county, state or country level to estimate a dose response function. Deschenes and Greenstone (2011) were the first to use the panel approach to quantify the impacts of climate change on residential electricity demand. They use variation residential energy consumption at the state level using flexible functional forms of daily average temperatures. Their identifying assumption is that wether fluctuations are random conditional on a set of spatial and time fixed effects, which is credible. The authors, like the time series papers cited above, find a U-shaped response function. 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. Aroonruengsawat and Auffhammer (2012) use a panel of household level electricity billing data to examine the impact of climate change on residential electricity consumption. They use within-household variation in temperature, which is made possible through variation in start dates and lengths of household billing periods. 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 (2011), ranging between 1% and 6%. The panel approach has the advantage that one can control for often extensive sets of fixed effects, which deal with the omitted variables issues the Ricardian model suffers from. This comes at a cost. The estimated response is again a short run - not a long run climate - response, which fails to incorporate extensive margin adaptation. Hsiang (2016) makes an argument that under a certain set of conditions, these short run responses are identical to climate responses, yet these are not necessarily met in the energy sector, where the penetration and efficiency of technology changes at a rapid pace. Further, the inclusion of large suites of fixed effects may amplify measurement error issues (Fisher et al., 2012).

A fourth approach, which to my knowledge has not yet been applied in the energy sector is long difference estimation. Burke and Emerick (2016) take long differences (e.g 10 or 20 year) of economic outcomes of interest (e.g. agricultural yields) and regress these on long differences of weather. This approach differences out until level unobservable cross sectional differences. The advantage of this method is that it estimates a long run climate response and is robust to the omitted variables issues raised in the Ricardian context. The data requirements are significant, as this approach requires a panel long enough to generate difference in weather, which is long enough to be interpreted as climate (ideally 30 years).

None of the approaches above formally model the intensive and extensive margin adaptation adjustments separately. The time series and panel approaches only capture intensive margin changes. The Ricardian and long difference estimation approaches have extensive margin adjustments "baked in". In order to separate the two one has to formally model the extensive margin adjustment process. A separate literature exists, which looks only at extensive margin adjustments, not necessarily just to higher temperatures but also income and prices. Biddle (2008) uses a reduced form econometric model, which accounts for changes in incomes, prices, and weather to explain the heterogeneity in air conditioner penetration throughout the 20th century. 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 ignoring the adoption decision would lead to an underestimation of future electricity consumption. Rapson (2011) estimates a dynamic structural model of air conditioner adoption using five cross sections of the Energy Information Administration's Residential Energy Consumption Survey (RECS), which could easily be expanded to estimate a weather/climate response.

Davis and Gertler (2015) is the only paper to my knowledge which combines a formal estimation of the extensive margin adoption decision with a more traditional panel data based intensive margin response function. They take advantage of a large database on household air conditioner penetration and electricity consumption for a large rapidly developing country - Mexico. They also employ a massive database of 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. The problem is that data on technology penetration and the outcome of interest (e.g. air conditioners and electricity consumption) do not exist for most developing or developed countries, which requires a different approach when one observes usage data only.

In this paper we propose a simple method which endogenizes the shape of the temperature response function without observing the level and type of adopted technology. This approach, CARE, uses fixed effects estimation to obtain causal estimates of the short run (intensive margin) temperature response for a large number of relatively fine spatial aggregates (ZIP codes). In a second regression we then estimate the sensitivity of the estimated slopes across space as a function of long run climate. The simulation then uses Global Climate Model (GCM) output to simulate the intensive margin impacts by moving along a given response function as well as the extensive margin impacts by shifting the response function as climate changes. This method is applicable beyond the energy sector. There is a literature which has endogenized response functions (e.g. (Barreca et al. 2016; Dell, Jones, and Olken 2012, 2014; Hsiang and Narita, 2012; Butler et al. 2013; Auffhammer and Aroonruengsawat (2012b)). We build on the general insight of a climate dependent response function and formalize an empirical approach to do so when one observes a large number of micro level observations on outcomes. This allows e.g. utilities to estimate the BAU impacts of climate change on consumption without having to engage in the costly collection of appliance stock and efficiency data.

The next section describes the data. Section (4) describes the empirical model, followed by estimation results in section (5). Section (6) contains the simulation results and section (7) concludes.

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 temporal data coverage for the four utilities by energy source (electricity and natural gas).

The dataset contains the complete bill level consumption and expenditure information for the population of single metered residential customers during the years we have data for as outlined in table (1). 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 or natural gas consumption (in kilowatt-hours, kWh or therms), 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. 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. 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. This is mostly not by choice of the homeowners, but simply due to the fact that not all of California has natural gas infrastructure to serve residences.

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. 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 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. Also, the bill start dates differ across households. Hereafter, this dataset is referred to as "billing data."

For electricity we observe total of 964 million bills and 928 million bills for natural gas. We observe 658 million electric bills for "normal" households, which are neither on the subsidized tariff nor all-electric homes. In addition we have 92 million bills for all-electric homes in the PG&E and SCE territory. The remaining bills are for households in all four utility territories on the subsidized tariff. We will treat these three classes of households separately in terms of estimation. For the simulation exercise we will take a consumption weighted average across these household types.

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 identify all ZIP codes across the three utilities' territories for which we have at least 1,000 bills. Figure 1 displays the approximately approximately 1,200 ZIP codes for which we observe such billing data. The left panel identifies the electricity ZIP codes and the right panel identifies the Natural Gas ZIP codes. Our sample covers a large portion of the state both from north to south and east to west. The sample represents approximately 80 percent of California's population.

3.2 Weather Data

The daily weather observations to be matched with the household consumption data have been provided by the PRISM (2004) project at Oregon State University. This dataset contains daily gridded maximum and minimum temperature for the continental United States at a grid cell resolution of roughly 2.5 miles. We observe these daily data for California from 1980-2015. In order to match the weather grids to ZIP codes, we have obtained a GIS layer of ZIP codes from ESRI, which is based on the US Postal Service delivery routes for 2013. For small ZIP codes not identified by the shape file we have purchased the location of these ZIP codes from a private vendor (zip-codes.com). We matched the PRISM 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 1980 - 2015.

3.3 Other Data

Unfortunately we 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 (2016).

There are 1,640 five digit ZIP codes, which have non-zero population in California. Our sample of ZIP does with more than 1,000 bills contains households for 1,165 of these. We do not have sufficient data for households in the remaining 475 ZIP codes. These remaining ZIP codes are either not served by the three utilities, or we do not have a sufficient number of bills 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 billing data. The ZIP codes in our sample represent 79 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 is a small but statistically significant difference in summer and winter temperature, without ZIP codes being slightly warmer. This is not surprising since most of the ZIP codes we are missing are in the Northern part of the state and the mountainous Sierras. The big difference in elevation confirms this. 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, population density, 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 estimated 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).

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

where $\log(q_{it})$ is the natural logarithm of household i's daily average electricity (or natural gas) consumed in kilowatt-hours (therms) 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 approximately 1,200 ZIP codes displayed in Figure 1. This is possible, since we observe such a large amount of data. While big data in past often posed a computational capacity problem once could overcome by sampling, it provides an opportunity in this context: the causal identification of a large number of electricity and natural gas 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. Obtaining ZIP code specific responses and simulation results will also allow us to examine the incidence of climate change on different socioeconomic groups in California exploiting cross sectional variation in population characteristics across ZIP codes. One could of course estimate a pooled regression with interaction terms to limit the number of estimated coefficients. This is simply a weighted average of our disaggregate results. Since one of the main points of this paper is the heterogeneity of impacts, we impose as little structure as possible by estimating equation (1) at the ZIP code level instead of pooling.

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, 2009; 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. 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 hotter bins, while the southwestern parts of California have few or no 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.[1] 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_{jp} 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_{jp}'s interpretation is approximately the percent increase in daily average household electricity/natural gas consumption during a billing period from experiencing one additional day in that temperature bin.

Panel (a) in Figure (2) displays the daily average temperature for the months of June, July and August averaged over the years 1981-2015. This is a reasonable measure of summer climate (a 25 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. Panel (b) displays the winter month (December, January, February) average daily temperature. The spatial distribution is similar to that of the summer climate. This figure simply stresses that due to its size and geography, California possesses significant heterogeneity in climate, which is necessary for our approach to work.

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. Further, we are not interested in identifying the price elasticity of demand here,

[1] The cutoffs for the bins are 24, 35, 40, 46, 51, 55, 59, 63, 67, 72, 78, 83 and 92 degrees Fahrenheit mean daily temperature.

since it is simply impossible write a better paper than Ito (2014), who uses the same electricity data we employ here.

An econometric issue arises that cannot be ignored. Higher temperatures in a given month will lead to higher electricity consumption. Block rate prices will force a share of households onto a higher pricing tier and raise average price, as is discussed in detail in Ito (2014). This induces a positive conditional correlation between price and consumption by design. This is testable as one would expect a positive coefficient estimate on price if included in model (1). Hence if we include price in equation (1) as part of Z_{it}, we would have to explicitly model the impact of higher temperatures on average price in a simulation framework. An alternate approach would be to omit average price from equation (1) and let the temperature coefficient capture both channels. If the intuition is correct here, what we would expect is that the pure temperature response is on average flatter in regressions that control for price. We will test for this and if confirmed omit average price from our regressions.

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. A third confounder, which we do not observe, is humidity. Humidity is not a major issue in California as most parts of the state are semi arid. Our temperature coefficients hence capture the effects of humidity. Our simulations would become invalid if the correlation patterns between humidity and temperature in the future were different from the historical correlations, which we could find no evidence of.

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) by fuel for each of the approximately 1,200 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 relatively short sample 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 Extensive Margin: 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 summer, the wealthy residents of San Francisco will install (additional) cooling equipment in their homes. To be clear - what we are interested in is the climate change driven response, not an income or price driven response. We provide an attempt to quantify the magnitude of this response. We estimate equations of the following form:

$$\beta_{jp} = \delta_1 + \delta_2 C_{pj} + \boldsymbol{\delta_3 Z_j} + \eta_j \tag{2}$$

where β_{jp} 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, which is why we limit the estimation of equation (2) to bins 10-14. 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, which is also a commonly used base temperature for calculating cooling degree days (CDD).

The variable C_{pj} in equation (2) is the share of days ZIP code j experienced in temperature bin p during the sample years 1981-2000 from our ZIP code level weather data produced from the PRISM data. C_{pj} 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 confounders 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). We also include population density to proxy for the level of urbanization. While we will not use the estimated coefficients on income and population density in our simulation later, controlling for them 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. We chose to estimate a pooled model, which restricts the coefficients on income and population density to be identical for all bins, yet controls for bin dummies. This provided more stable estimation results than the separate equations. The bin dummies control for the fact that each bin contains the response coefficients for a different collection of ZIP codes. This arises, as we mentioned above, since we do not have temperature coverage in all bins for all ZIP codes. Finally, we estimate equation (2) via Ordinary Least Squares with heteroskedasticity robust standard errors, as the dependent variables are estimated coefficients and do not have constant variance. Running Weighted Least Squares does not significantly change the results, yet the least squares estimates are more stable.

5. Estimation Results

5.1 Intensive Margin: The Usage Response to Temperature

As discussed in the previous section, we estimate equation (1) for each of the approximately 1,200 ZIP codes that have more than 1,000 bills. While we cannot feasibly present the approximately 1,200 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 top figure estimates equation (1) by excluding average price from Z_{it}. The thick black line displays the median temperature response curve across the approximately 1,200 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 trough 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. The bottom panel displays the temperature response when including average price in the regressions. As hypothesized, the response function flattens out significantly after controlling for price, which is consistent with the forced positive correlation between average price and consumption.

Figure (4) produces analogous pictures for the subsidized households and the all-electric households. The subsidized household 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.

Figure (5) displays the analogous results for the natural gas regressions, also excluding price from the regressions. Since the only major ambient temperature sensitive use of natural gas in residences is space heating, we would expect a downward sloping line in temperature at low temperatures and a relatively flat response curve at higher temperatures. Figure (5) displays exactly that. There is quite a bit of variation in slope across ZIP codes, yet the median response is exactly as expected and flattens out at almost exactly 65 degrees Fahrenheit.

5.2 Extensive Margin: The Long-Run Electricity Consumption Response to Temperature

As discussed in Section 4.2, we exploit the approximately 1,200 estimated electricity temperature response curves and examine whether we can explain variation in temperature response at high temperatures through cross-sectional variation in "climate" as well as income and population density.

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 during the years 1981-2000 (our predetermined proxy for summer climate), income and population density. We run a pooled regression, the results of which are shown in table (3). As the dependent variable is an estimated coefficient, we use White robust standard errors. Model (1) simply regressed the slope for the roughly 5,000 ZIP/bin observations on the share of days spent in the bin. The coefficient is negative, which hints at omitted variables issues. In model (2) we control for income and population density 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 theoretically correct positive sign, which indicates that having more days in the bins, which usually see some cooling, results in a steeper temperature response. This is not surprising as the bins contain very different mixes of ZIP codes. Not all ZIP codes, for example, experience days in the hottest bin. The bin fixed effects hence control for unobservable difference across bins. 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. Model (5) is identical to Model (4) but does the regression for the subsidized households. Model (6) conducts the regression for all electric homes.

We use the results from models (4), (5) and (6) 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 and Natural Gas Consumption Simulations

In this section, we simulate the impacts of climate change on electricity and then natural gas 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. For electricity 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 from the previous section. 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.

In our simulations we make one key assumption. For natural gas we only use the intensive margin simulations as one would not expect households to install more or fewer heaters in response to climate change. We would expect existing equipment to be operated less frequently. But one would not go an install a more efficient heater which is going to be used less due to climate change.

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 (Taylor et al. 2012) utilizing a modification (Hegewisch and Abatzoglou, 2015) of the Multivariate Adaptive Constructed Analogs (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.[2]

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 (6) displays the projected increases in household residential electricity consumption across the approximately 1,200 ZIP codes for each of the four projection periods and the 18 GCMs for RCP8.5. 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.[3]. 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. Panel (a) in Figure (7) 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 projections ignore potential extensive margin impacts, which we turn to next.

For each ZIP code, climate model and scenario, we calculate the simulated shift of 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 (8) collapses the temperature response

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

[3] We trim the distribution of estimated impacts at the top and bottom as some point estimates are too large to be credible. This has to do with a lack of precision for some zip codes with very few observations in the extreme bins. We censor the slope coefficients to be less than 0.2 in absolute value and projected impacts to be less than 50%.

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 (6) displays the same box and whisker plots as we did for the intensive margin simulations earlier, but now incorporating the extensive margin changes. 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.

Panel (b) in 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.= fro the extensive margin adaptation. It indicates a noticeable increase in consumption across the state. The right panel 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 increases the estimated impacts by 50%. The second noteworthy fact is that even until 2059, the estimated impacts for electricity consumption are relatively small - strictly less than 5% even for the worst case scenario. In terms of electricity planners planning horizon 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.

For natural gas, however, we see more significant decreases in consumption, even by mid-century. Under RCP8.5 consumption is expected to decrease by 10.4% and by end of century by 20.5%. While again, the end of century is a long ways away and beyond the utility planners horizon, this begs the question whether in this counterfactual world, the savings from natural gas are bigger than the projected increases in electricity consumption. The EIA states that California Homes used 0.287 quadrillion BTU of electricity and 0.439 quadrillion BTU of natural gas in 2009. If we use the projected percentage changes from table (4) we arrive at the conclusion that climate change is simulated to lead to a 0.039 quad BTU net decrease in energy consumption for the residential sector in California. We will discuss the limitations of this simulation in the conclusions, yet before we do it is instructive to put into perspective the impacts of other drivers for electricity consumption over the next century.

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 27 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.Natural gas demand would of course increase as well as more people will demand more gas.

6.3 Incidence of Climate Change

There is big literature on the sorting of individuals across space in order to match amenities to their preferences. Local climate is of course one of the characteristics that individuals take into account when choosing a place to live. There is an emerging literature suggesting that climate shocks may lead to measurable rural to urban migration (e.g. Feng et al. 2010, Auffhammer and Vincent 2012 and Feng et al. 2012). We would expect that shifts in climate would lead to a redistribution of population across space. Transactions costs of moving are large, and it would likely require significant climate change for there to be a detectable redistribution of population.

It is therefore instructive to examine whether there is a correlation between projected impacts of electricity consumption and current observable population characteristics. Table (6) provides such conditional correlations. We regress the projected total impacts across all climate models for mid and end of century and RCPs 4.5 and 8.5 on ZIP code observables. These clearly non-causal estimates do provide some interesting patterns. It seems that ZIP codes with less expensive homes, a larger share of hispanics, higher income and a younger population are projected to experience higher impacts. This is not surprising, given that Figure (7) shows that the main impacts will be concentrated in the San Joaquin Valley and the interior parts of Southern California. The sign on income is somewhat surprising, yet it is conditional on home value.

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 demand 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.

By estimating the response of natural gas consumption to higher temperatures, we show that the projected increases in electricity consumption are more than offset by savings in natural gas, making climate change a net energy saving factors for the residential sector. It is important to keep in mind several caveats. These are not forecasts, yet simulations. We think of the results provided in this paper as imposing end of century climate on a current day economy. Many other drivers of energy consumption will change. What our paper shows is the business as usual path, which mitigation strategy has to work against. 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 electricity consumption outlined in this paper and amplify the natural gas savings. 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] Abatzoglou, John T., and Timothy J. Brown. 2012. "A comparison of statistical downscaling methods suited for wildfire applications." International Journal of Climatology 32.5: 772-780.

[2] Albouy, D., Graf, W. F., Kellogg, R., & Wolff, H. 2013. Climate amenities, climate change, and American quality of life (No. w18925). *National Bureau of Economic Research.*

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

[4] Auffhammer, Maximilian and Anin Aroonruengsawat. 2012b Hotspots of climate-driven increases in residential electricity demand: A simulation exercise based on household level billing data for California. California Energy Commission.

[5] Auffhammer, Maximilian, Patrick Baylis, and Catherine H. Hausman. 2017. "Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States." Proceedings of the National Academy of Sciences: 201613193.

[6] Auffhammer, Maximilian, and Erin T. Mansur. 2014. "Measuring climatic impacts on energy consumption: A review of the empirical literature." Energy Economics 46: 522-530.

[7] Auffhammer, M. and Vincent, J. R. 2012. Unobserved time effects confound the identification of climate change impacts. Proceedings of the National Academy of Sciences, 109(30), 11973-11974.

[8] 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

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

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

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

[12] Burke, Marshall and Kyle Emerick. 2016. "Adaptation to climate change: Evidence from US agriculture". American Economic Journal -Economic Policy, 8(3), 106-140.

[13] Butler, Ethan E., and Peter Huybers. 2013. "Adaptation of US maize to temperature variations." Nature Climate Change 3.1: 68-72.

[14] Carleton, Tamma A., and Solomon M. Hsiang. 2016. "Social and economic impacts of climate." Science 353.6304 : aad9837.

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

[16] 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.

[17] 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

[18] Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. "Temperature shocks and economic growth: Evidence from the last half century." American Economic Journal: Macroeconomics 4.3: 66-95.

[19] Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What do we learn from the weather? The new climateeconomy literature." Journal of Economic Literature 52.3: 740-798.

[20] 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.

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

[22] 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.

[23] Feng, S., Oppenheimer, M., and Schlenker, W. 2012. Climate change, crop yields, and internal migration in the United States (No. w17734). National Bureau of Economic Research.

[24] Fisher, Anthony C., et al. 2012. "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather: comment." The American Economic Review 102.7: 3749-3760.

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

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

[27] Hsiang, Solomon M., and Daiju Narita. 2012. "Adaptation to cyclone risk: Evidence from the global cross-section." Climate Change Economics 3.02: 1250011.

[28] Hsiang, Solomon. 2016. "Climate econometrics." Annual Review of Resource Economics 8: 43-75.

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

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

[31] 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.

[32] Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., ... and Lettenmaier, D. P. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. Journal of Climate, 26(23), 9384-9392.

[33] 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.

[34] 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.

[35] Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. 1994.“The impact of global warming on agriculture: a Ricardian analysis.” The American economic review : 753-771.

[36] Feng, S., Krueger, A. B. and Oppenheimer, M. 2010. Linkages among climate change, crop yields and MexicoUS cross-border migration. Proceedings of the National Academy of Sciences, 107(32), 14257-14262.

[37] Pierce, David W., et al. 2015. “Improved bias correction techniques for hydrological simulations of climate change.” Journal of Hydrometeorology 16.6: 2421-2442.

[38] PRISM Climate Group, Oregon State University. 2004. .

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

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

[41] Rose, Steven, D. Turner, G. Blanford, J. Bistline, F. de la Chesnaye, and T. Wilson. 2014. “Understanding the social cost of carbon: A technical assessment.” EPRI Technical Update Report.

[42] 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.

[43] 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.

[44] 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.

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

[46] 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.

[47] Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl. 2012. "An overview of CMIP5 and the experiment design." Bulletin of the American Meteorological Society 93.4: 485-498.

[48] 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

Notes: This table displays the total number of bills in our dataset. We drop electricity bills with average daily consumption less than 2kWh as well as solar homes. Further, our models only include ZIP codes for which we have more than 1,000 bills.

Table 2: Summary Statistics for ZIP Codes In and Out of Sample

	In Sample	Out of Sample	p-value
Population (in thousands)	25.19	16.51	0.00
% White	70.07	72.35	0.04
% Black	5.13	5.31	0.67
% Hispanic	30.95	26.08	0.00
% Asian	10.87	10.10	0.26
% Male	50.14	50.93	0.00
Median Age (years)	38.90	40.31	0.00
Persons per Household	2.85	2.59	0.00
Average Home Value (in 100k US$)	4.14	3.98	0.39
Income per Household (in 10k US$)	6.52	5.99	0.00
Population Density	30.21	44.51	0.00
Elevation (in feet)	392.10	741.26	0.00
Mean Summer Temperature (F)	72.03	70.51	0.00
Mean Winter Temperature (F)	50.74	48.48	0.00
Mean Summer Precipitation (mm)	0.10	0.16	0.00
Mean Winter Precipitation (mm)	3.25	3.43	0.13

Notes: This table displays the mean observable characteristics of the ZIP codes in our sample and ZIP codes not in our sample with positive population. The t-test assumes unequal variances. The observable characteristics were purchased from zip-codes.com.

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

	(1)	(2)	(3)	(4)	(5)	(6)
Historical Bin	-0.047***	-0.0450***	0.0124***	0.0116***	0.0217***	0.0168***
Tavg Share	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Interaction				0.0276***	0.0758***	0.0858***
Bin 12+				(0.000)	(0.000)	(0.000)
Special Customer	No	No	No	No	Care	All-E
Income	No	Yes	Yes	Yes	Yes	Yes
Population Density	No	Yes	Yes	Yes	Yes	Yes
Bin Fixed Effects	No	No	Yes	Yes	Yes	Yes
Price in Equation (1)?	No	No	No	No	No	No
Observations	5,137	5,116	5,116	5,116	4,984	4,642

Notes: This table displays coefficients from a regression of the electricity slope coefficients estimated in equation (1) on the share of days in a given temperature bin the ZIP code has experienced over the period 1981-2000. The regression only includes the air conditioning relevant temperature bins 10-14. The standard errors are Huber-White. Regressions 1-4 are for "normal" households. Regression (5) is for households with subsidized energy bills. Regression (6) is for all-electric homes.

Table 4: Projected Percent Changes in Residential Electricity Consumption

Simulation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
RCP	4.5	8.5	4.5	8.5	4.5	8.5	4.5	8.5	4.5	8.5
Special Customer	No	No	No	No	CARE	CARE	All-E	All-E	No	No
Fuel	Elec.	Elec.	Elec.	Elec.	Elec.	Elec.	Elec.	Elec.	Gas	Gas
Price Controls	No	No	Yes	Yes	No	No	No	No	No	No
Intensive Margin										
2020-39	0.8	1.1	0.5	0.7	0.8	1.1	-0.2	-0.2	-4.0	-4.9
2040-59	2.2	3.2	1.5	2.3	2.0	2.9	0.0	0.3	-7.9	-10.4
2060-79	3.2	6.7	2.3	4.7	2.9	6.0	0.3	1.9	-10.3	-16.1
2080-99	3.7	10.8	2.6	7.4	3.3	9.8	0.5	4.3	-11.3	-20.5
Extensive Margin										
2020-39	1.2	1.6	1.1	1.4	1.1	1.5	0.3	0.4	NA	NA
2040-59	3.2	4.8	2.9	4.4	3.0	4.5	1.2	1.9	NA	NA
2060-79	4.8	10.6	4.3	9.6	4.5	10.1	1.9	5.4	NA	NA
2080-99	5.6	17.6	5.1	15.9	5.3	16.9	2.4	10.2	NA	NA

Notes: This table displays the simulated percent increase in total residential electricity consumption relative to 2000-2015 climate for the two IPCC Representative Concentration Pathways with low emissions (4.5) and high emissions (8.5). Columns 1 and 2, indicate simulated increases for normal households without controlling for price in the regressions. Columns (3) and (4) control for price. Columns (5) and (6) simulate increases for subsidized households. Columns (7) and (8) simulate changes for households which are all-electric. Columns (9) and (10) display the impacts on natural gas consumption for households with gas bills.

Table 5: 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	17%	19%	66%	68%	103%	104%
2060-79	21%	26%	93%	99%	171%	176%
2080-99	27%	39%	122%	134%	275%	287%

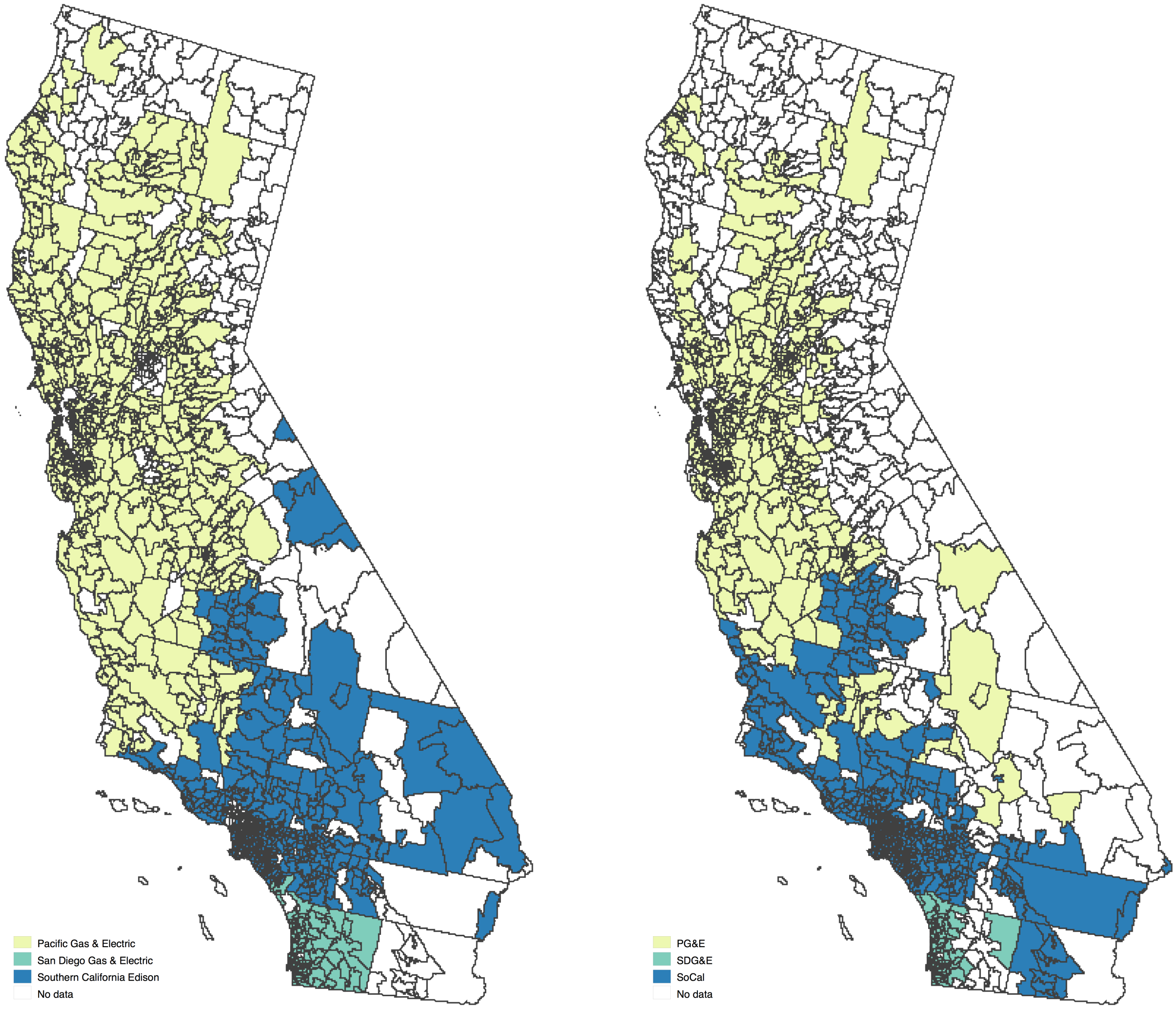
Note: This table displays the simulated percent increase in total residential electricity consumption relative to 2010 population and 2006-2015 climate for low, medium and high population growth scenarios using both intensive and extensive margin adjustments. The figures are population weighted averages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

Table 6: Correlations between Impacts and Population Characteristics

Outcome: Projected Impacts in %	(1)	(2)	(3)	(4)
% White	0.01	0.00	0.01	0.01
	(0.008)	(0.028)	(0.006)	(0.010)
% Black	-0.01	-0.01	-0.00	-0.01
	(0.012)	(0.038)	(0.008)	(0.013)
% Hispanic	0.03***	0.09***	0.02***	0.03***
	(0.006)	(0.021)	(0.004)	(0.007)
Average Home Value (in 100k$)	-0.50***	-1.40***	-0.33***	-0.56***
	(0.046)	(0.142)	(0.032)	(0.052)
Household Income (in 10k$)	0.49***	1.58***	0.33***	0.56***
	(0.049)	(0.157)	(0.034)	(0.056)
Median Age	-0.06***	-0.20***	-0.04***	-0.07***
	(0.020)	(0.062)	(0.013)	(0.022)
Constant	4.08***	15.37***	2.60***	4.91***
	(1.071)	(3.470)	(0.733)	(1.212)
RCP	8.5	8.5	8.5	8.5
Period	2040-59	2080-99	2040-59	2080-99
Observations	1,165	1,165	1,165	1,165
R-squared	0.209	0.194	0.203	0.208

Note: This table displays simple and very much non-causal regressions of ensemble average predicted impacts including intensive and extensive margin effects on a number of observable characteristics of the population across ZIP codes.

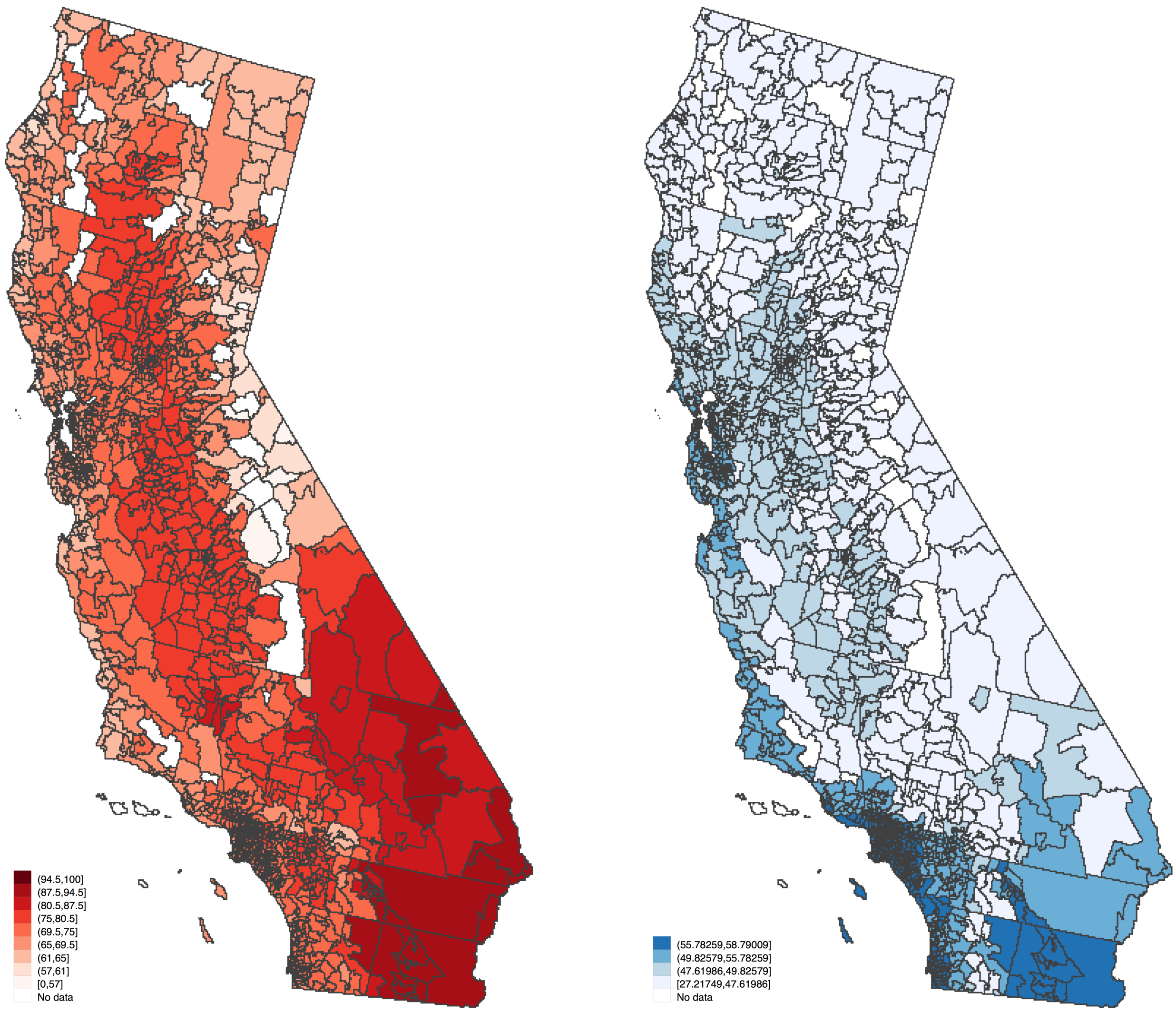
Figure 1: ZIP Codes with observed residential electricity and Natural Gas Bills by Investor Owned Utility.



(a) Electricity Sample (b) Natural Gas Sample

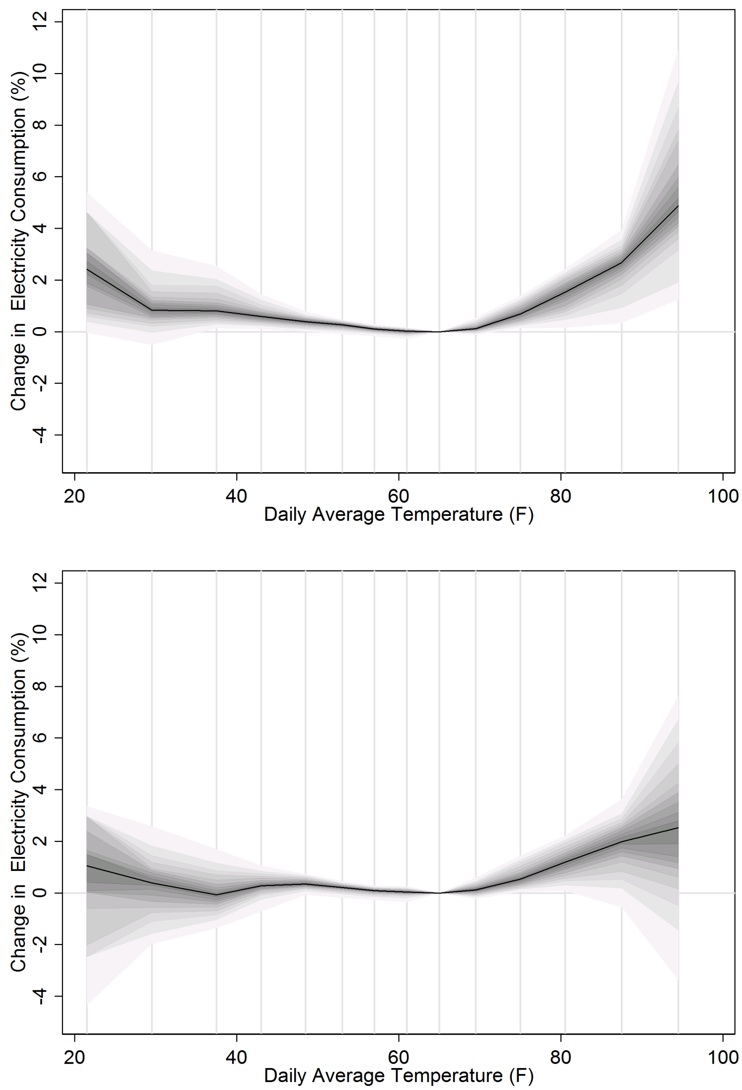
Notes: The map above displays the five-digit ZIP codes for which we have more than 1000 bills over the estimation period from either PG&E, SCE, SDG&E or SoCalGas. Zip codes with no data either have fewer than 1000 bills total or a reserved by one of California's many municipal utilities. Due to the small size of many ZIP codes they do not show up in the map at the current resolution.

Figure 2: California's Summer (June-August) and Winter (December-February) Climate: Average Daily Temperature 1981-2015



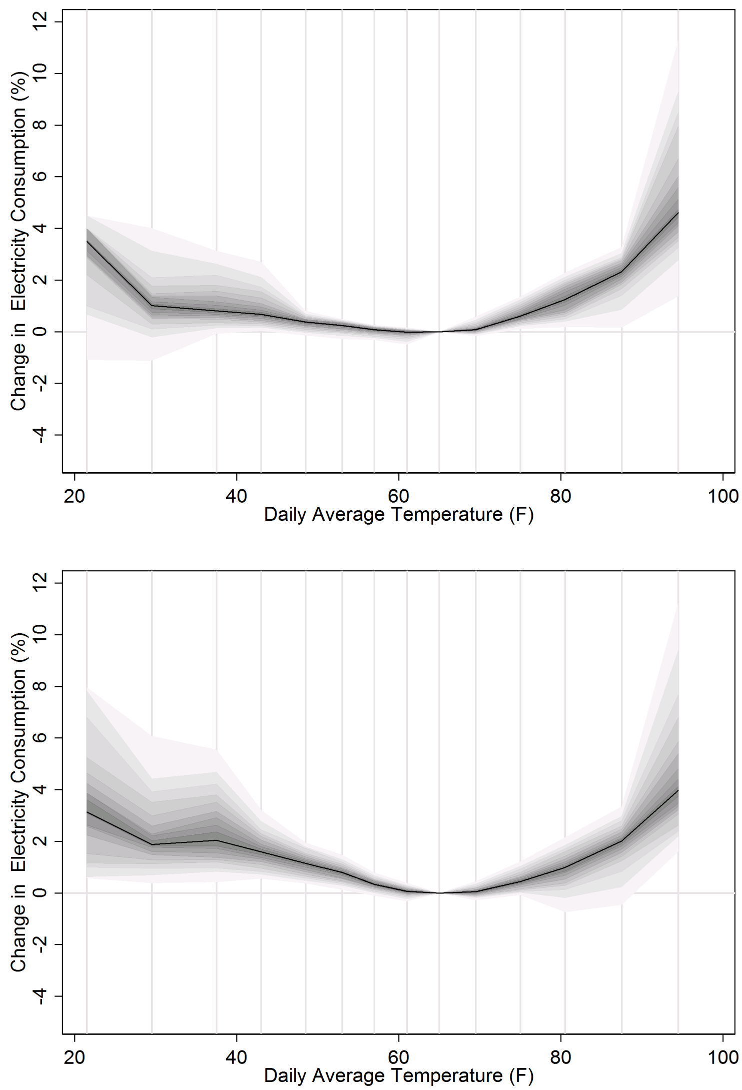
(a) Summer (b) Winter

Figure 3: Distribution of Estimated Electricity Temperature Response Coefficients across ZIP Codes omitting price (top panel) and controlling for price (bottom panel)



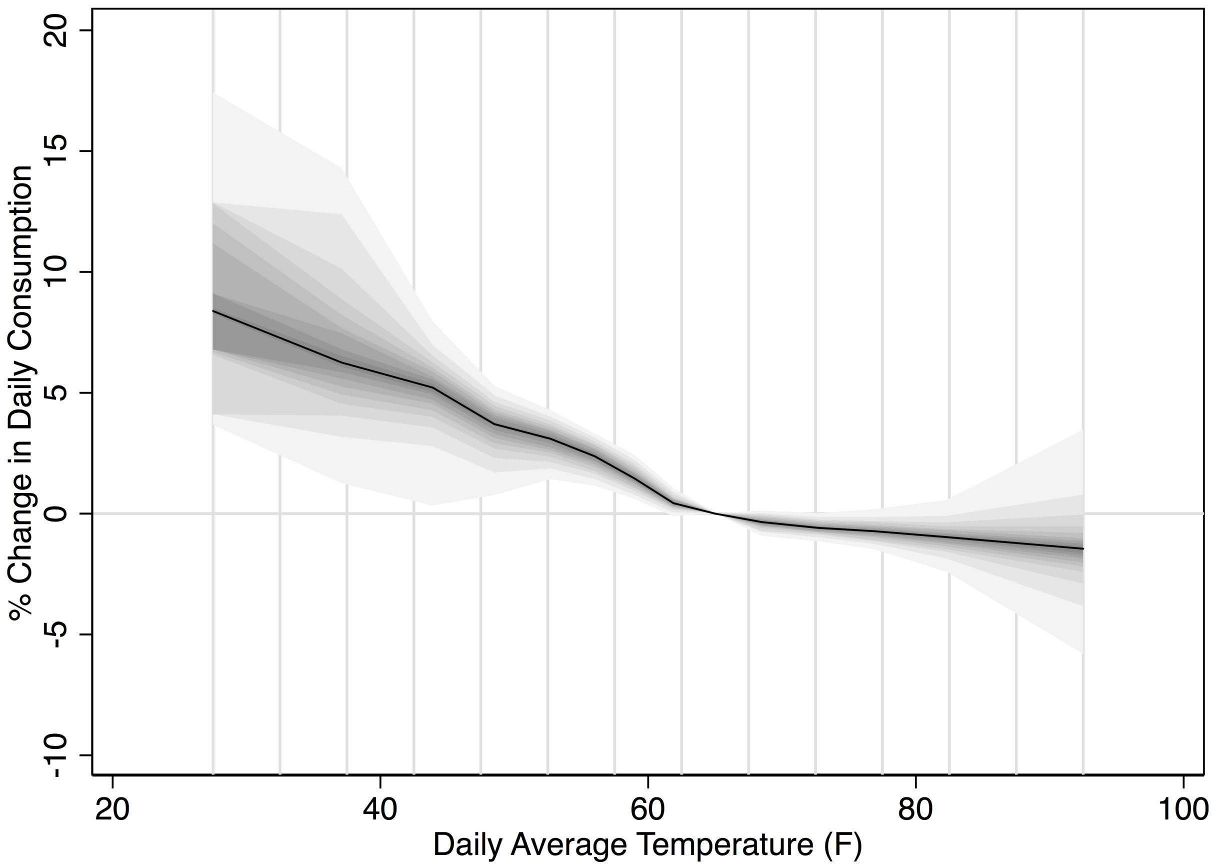
Notes: This figure displays the empirical distribution of the estimated electricity temperature response function across ZIP codes in the sample across percentile temperature bins. The lightest gray shading indicates the range of the 5^{th} to 95^{th} percentile. Each darker shading represents a 5% increase in the percentile. The solid black line represents the median temperature responsiveness. The vertical grey lines indicate the cutoffs of the temperature bins. The top panel displays the distribution of the response functions without controlling for average price. The bottom panel displays the distribution after controlling for average price.

Figure 4: Distribution of Estimated Electricity Temperature Response Coefficients across ZIP Codes for subsidized households (top panel) and all electric households (bottom panel)



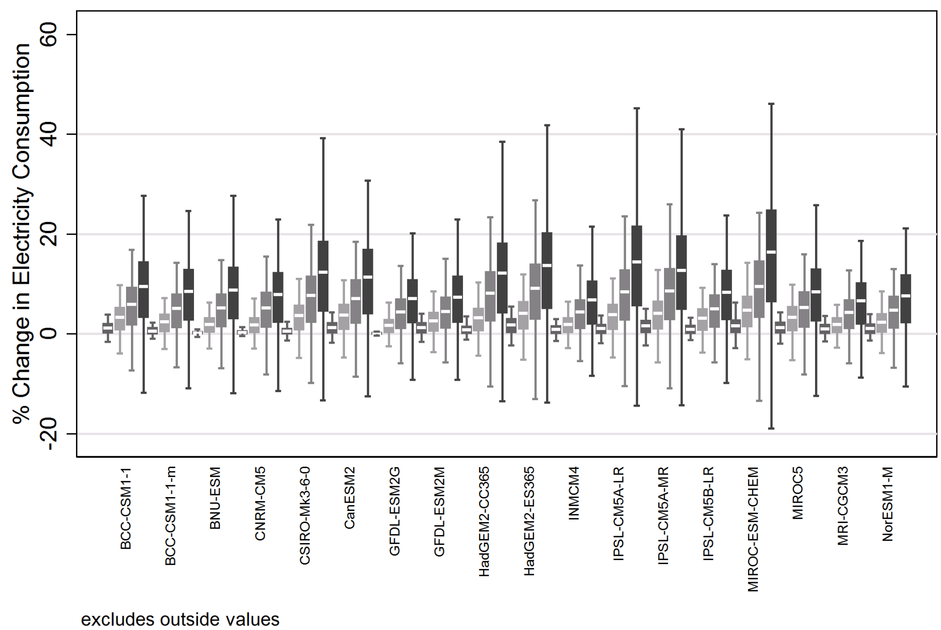
Notes: This figure displays the empirical distribution of the estimated electricity temperature response function across ZIP codes in the sample across percentile temperature bins. The lightest gray shading indicates the range of the 5^{th} to 95^{th} percentile. Each darker shading represents a 5% increase in the percentile. The solid black line represents the median temperature responsiveness. The vertical grey lines indicate the cutoffs of the temperature bins. The top panel displays the distribution of the response functions for households who receive a discount on their electric bill under the subsidy program (California Alternate Rates for Energy). The bottom panel displays the distribution for all-electric households. Neither panel controls for price in the underlying regressions.

Figure 5: Distribution of Estimated Natural Gas Temperature Response Coefficients across ZIP Codes



Notes: This figure displays the empirical distribution of the estimated natural gas 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. The regressions do not control for average price.

Figure 6: Intensive margin [Top Panel] and Extensive Margin [Bottom Panel] per household impacts across ZIP codes and Climate Models.



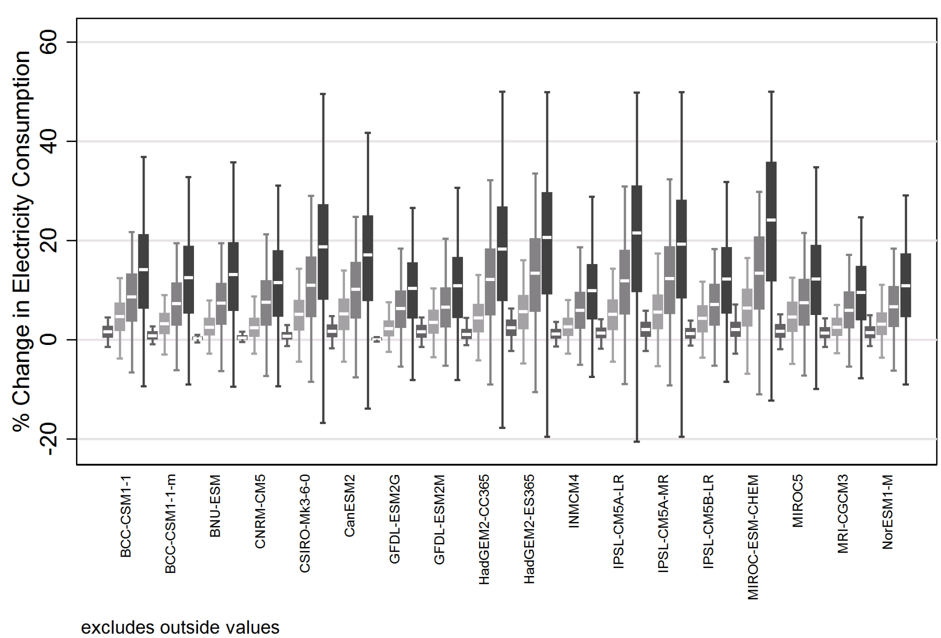
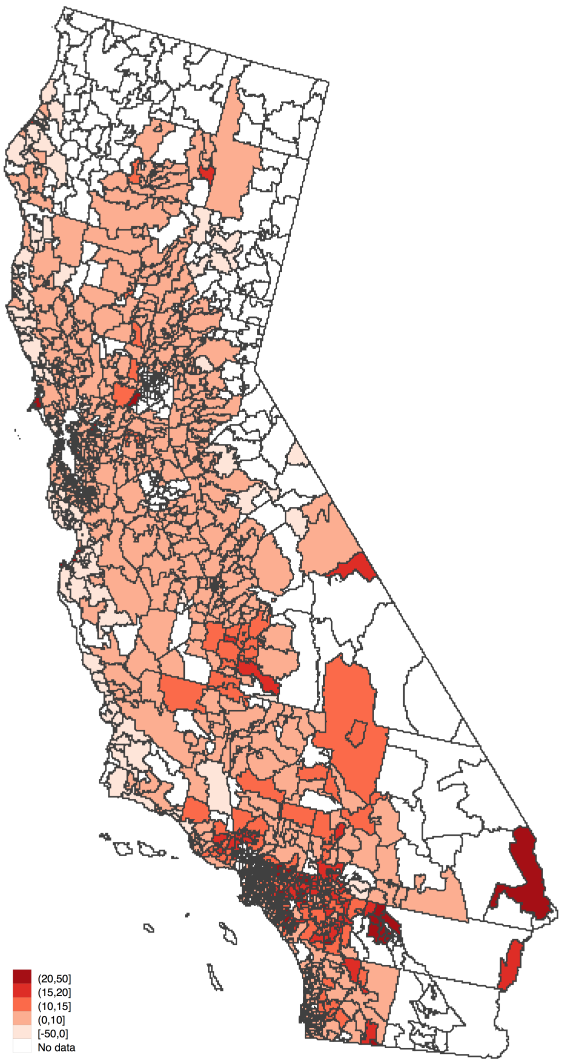
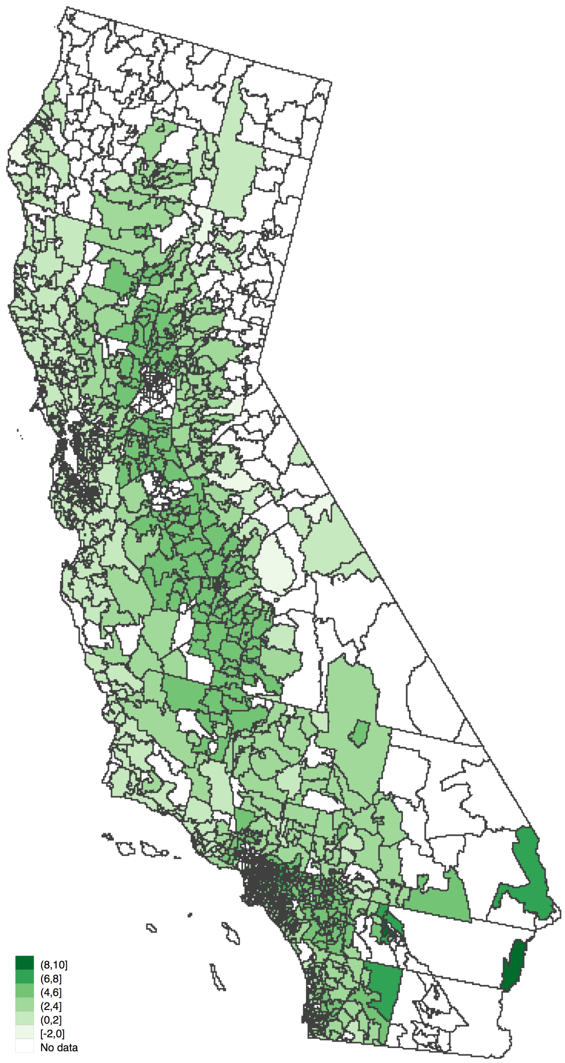


Figure 7: Intensive and Extensive Margin Adjustment: Projected Percent Increases in Average Household Electricity Consumption 2080-2099 over 2000-2015 for RCP 8.5



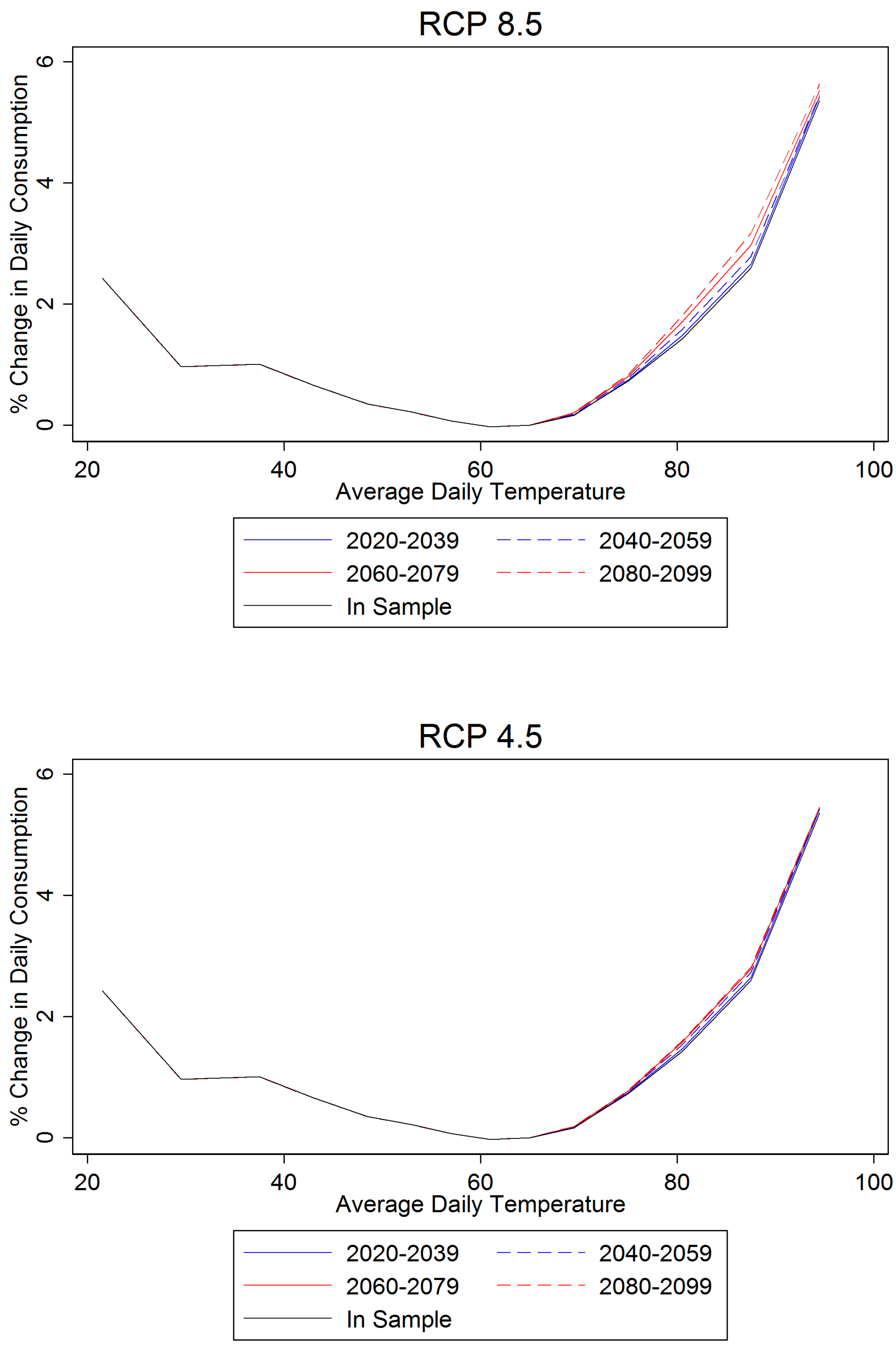
(a) Intensive Margin



(b) Extensive Margin Delta

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 8: 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.