Title
Essays on the Impact of Climate Change and Building Codes on Energy Consumption and the Impact of Ozone on Crop Yield

Author
Aroonruengsawat, Anin

Publication Date
2010

Essays on the Impact of Climate Change and Building Codes on Energy Consumption and the Impact of Ozone on Crop Yield

by

Anin Aroonruengsawat

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Maximilian Auffhammer, Chair
Professor Michael Hanemann
Professor Catherine Wolfram

Spring 2010

Essays on the Impact of Climate Change and Building Codes on Energy Consumption and the Impact of Ozone on Crop Yield

Abstract

Essays on the Impact of Climate Change and Building Codes on Energy Consumption and the Impact of Ozone on Crop Yield

by

Anin Aroonruengsawat
Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Maximilian Auffhammer, Chair

Part I is a joint paper with Maximilian Auffhammer. The study simulates the impacts of higher temperatures resulting from anthropogenic climate change on residential electricity consumption for California. Flexible temperature response functions are estimated by climate zone, which allow for differential effects of days in different temperature bins on households' electricity consumption. The estimation uses a comprehensive household level dataset of billing data for California's three investor-owned utilities (Pacific Gas and Electric, San Diego Gas and Electric, and Southern California Edison). The results suggest that the temperature response varies greatly across climate zones. Simulation results using a downscaled version of the National Center for Atmospheric Research global circulation model suggest that holding population constant, total consumption for the households considered may increase by up to 55% by the end of the century.

Part II is a joint work with Maximilian Auffhammer and Alan Sanstad. We study the impacts of state level residential building codes on per capita residential electricity consumption. We construct a timeline of when individual states first implemented residential building codes. Using panel data for 48 US states from 1970-2006, we exploit the temporal and spatial variation of building code implementation and issuance of building permits to identify the effect of the regulation on residential electricity consumption. Controlling for the effect of prices, income, and weather, we show that states that adopted building codes followed by a significant amount of new construction have experienced detectable decreases in per capita residential electricity consumption - ranging from 3-5% in the year 2006. Allowing for heterogeneity in enforcement and code stringency results in larger estimated effects.

In the last part, I estimate the impact of ground level ozone on corn and soybean yields using nation-wide county-level data and ozone measures for the U.S. during 1990-2006. The implementation of the NOx Budget Trading Program (NBP) aiming to reduce NOx and thus ozone during the growing season is used as an instrument to control for endogeneity in the yield regression. The estimated elasticites of soybean

and corn yield with respect to seasonal mean ozone concentrations are -0.60 and -0.57 respectively. The estimated elasticities of crop yield with respect to maximum ozone concentrations are higher suggesting a nonlinear relationship. A back of the envelope calculation shows that soybeans and corn loss from a one standard deviation increase in mean ozone during growing season is about $5 billion dollars. The NBP program reduces the value of crop losses by 2.19 billion for soybeans and 3.2 billion dollars corn during 2003 to 2007.

For my parents.

Acknowledgments

The completion of this dissertation would not have been possible without my amazing advisor Maximilian Auffhammer. He has given me invaluable advices and has been tremendously patient with me. I really appreciate what you have done for me. THANK YOU. I also would like to thank Michael Hanemann and Catherine Wolfram for their comments to improve my dissertation.

During the Ph.D. quest, I have met many people who become my best friends. In Atlanta, I am lucky to meet Thalyta Yuwono who always cheer me up when I am down. Life in Berkeley would have been bored without Quoc Luong, Damian Bickett, Hideyuki Nakagawa and Sarah Dobson. I thank Quoc for his friendship and various kind of supports especially some extra cash. I thank Damian for organizing an amazing trip to Zion. It is so far the best trip ever in my life. I thank Hide for accompanying me for a coffee break and lunch. I thank Sarah for free ”‘venting-session”’ and for slowing down for me on our bike ride. I will miss you guys so much. I also would like to thank Nitinant Wisaweisuan who always pray for me from the other side of the world.

Last but the most important, I thank my parents for raising me and for being patient waiting for me. The day has come and your son is now a Ph.D. This is for you.

Contents

List of Figures

List of Tables

Part I

Impacts of Climate Change on Residential Electricity Consumption: Evidence from Billing Data

Chapter 1

Introduction

Forecasts of electricity demand are of central importance to policy makers and utilities for purposes of adequately planning future investments in new generating capacity. Total electricity consumption in California has more than quadrupled since 1960, and the share of residential consumption has grown from 26% to 34% (EIA SEDS 2008). Today, California's residential sector alone consumes as much electricity as Argentina, Finland, or roughly half of Mexico. The majority of electricity in California is delivered by three investor-owned utilities and over a hundred municipal utilities.

On a per capita basis, California's residential consumption has stayed almost constant since the early 1970s, while most other states have experienced rapid growth in per capita consumption. The slowdown in growth of California's per capita consumption coincides with the imposition of aggressive energy efficiency and conservation programs during the early 1970s. The average annual growth rate in per capita consumption during 1960-1973 was approximately 7% and slowed to a remarkable 0.29% during 1974-1995. Growth rates during the last decade of available data have increased to a higher rate of 0.63%, and this difference in growth rates is statistically significant.

California's energy system faces several challenges in attempting to meet future demand (CEC 2005). In addition to rapid population growth, economic growth and an uncertain regulatory environment, the threat of significant global climate change has recently emerged as a factor influencing the long term planning of electricity supply. The electric power sector will be affected by climate change through higher cooling demand, lower heating demand, and potentially stringent regulations designed to curb emissions from the sector.

This paper simulates how the residential sector's electricity consumption will be affected by different scenarios of climate change. We make three specific contributions to the literature on simulating the impacts of climate change on residential electricity consumption. First, through an unprecedented opportunity to access the complete billing data of California's three major investor-owned utilities, we are able to provide empirical estimates of the temperature responsiveness of electricity consumption based on micro-data. Second, we allow for a geographically specific response of electricity consumption to changes in weather. Finally, we explore socio-economic and physical characteristics of the population, which help explain some of the variation in temperature response.

The paper is organized as follows: Chapter 2 reviews the literature assessing the impacts of climate change on electricity consumption. Chapter 3 describes the sources of the data used in this study. Chapter 4 contains the econometric model and estimation results. We simulate the impacts of climate change on residential electricity consumption in chapter 5. Chapter 6 explores the heterogeneity in temperature response and chapter 7 concludes.

Chapter 2

Literature Review

The historical focus of the literature forecasting electricity demand has been on the role of changing technology, prices, income, and population growth (e.g., Fisher and Kaysen 1962). Early studies in demand estimation have acknowledged the importance of weather in electricity demand and explicitly controlled for it to prevent biased coefficient estimates, as well as wanting to gain estimation efficiency (e.g., Houthakker and Taylor, 1970). Simulations based on econometrically estimated demand functions had therefore focused on different price, income, and population scenarios, while assuming a stationary climate system. The onset of anthropogenic climate change has added a new and important dimension of uncertainty over future demand, which has spawned a small academic literature on climate change impacts estimation, which can be divided 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 number of specific and often untestable assumptions about the evolution of the capital stock and its usage. 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 in his seminal book *The Economics of Climate Change*. The section dealing with the impact on space cooling and heating 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 and finds that increases in annual temperatures ranging from 1.0°C-1.4°C (1.8°F-2.5°F) in 2010 would result in demand of 9% to 19% above estimated new capacity requirements (peak load and base load) in the absence of climate change. The estimated impacts rise to 14% and 23% for the year 2055 and an estimated 3.7°C (6.7°F) temperature increase.

Baxter and Calandri (1992) provide another early study in this literature and focus on California's electricity use. In their study they utilize a partial equilibrium model of the residential, commercial, agriculture, and water pumping sectors, to examine total consumption as well as peak demand. They project electricity demand for these sectors to the year 2010 under two global warming scenarios: a rise in average annual temperature of 0.6°C (1.1°F) (Low scenario) and of 1.9°C (3.4°F) (High scenario). They find that electricity use increases from the constant climate scenario by 0.6% to 2.6%, while peak demand increases from the baseline scenario by 1.8% to 3.7%. Rosenthal et al. (1995) focus on the impact of global warming on energy expenditures for space heating and cooling in residential and commercial buildings. They estimate that a 1°C (1.8°F) increase in temperature will reduce U.S. energy expenditures in 2010 by $5.5 billion (1991 dollars).

The economics literature has favored the 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.

Instead of looking at a cross section of firms or households, 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 (CalISO) for 2004 and regress it on a population weighted average of 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 global circulation models (GCMs) forced using three Intergovernmental Panel for Climate Change (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%-20.3% and 1.0%-19.3%, 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% of actual consumption, which is similar to the impact estimated by Baxter and Calandri (1992).

Deschênes and Greenstone (2007) provide the first panel data-based approach to estimating the impacts of climate change on residential total energy consumption, which includes electricity, natural gas and oil as the main non0renewable sources of energy. 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 consumption. 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 electricity consumption is higher on very cold and hot days. The impact of climate change on annual electricity consumption by 2099 is in the range of 15%-30% of the baseline estimation or 15 to 35 billion (2006 US$). 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.

The current paper is the first paper using a panel of household level electricity billing data to examine the impact of climate change on residential electricity consumption. Through a unique agreement with California's three largest investor-owned utilities, we gained access to their complete billing data for the years 2003-2006. We identify the effect of temperature on electricity consumption using within household variation in temperature, which is made possible through variation in the start dates and lengths of billing periods across households. Since our dataset is a panel, we can control for household fixed effects, month fixed effects, and year fixed effects. The drawback of this dataset is that the only other reliable information we have about each individual household is price and the five-digit ZIP code location.

Chapter 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 Pacific Gas and Electric, Southern California Edison, and San Diego Gas and Electric for the years 2003-2006. These three utilities provide electricity to roughly 80% of California households.

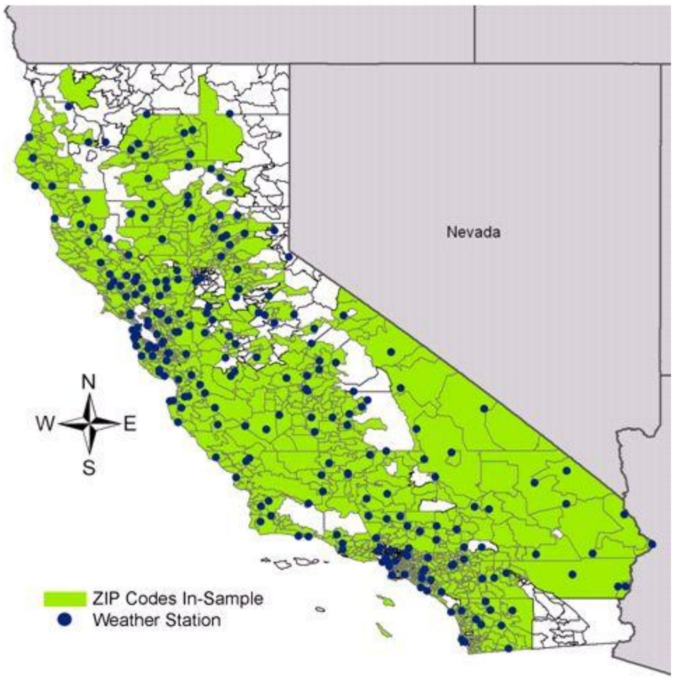
The data set contains the complete information for each residential customer's bills over the four year period. Specifically, we observe an ID for the physical location, a service account number[1], bill start-date, bill end-date, total electricity consumption (in kilowatt-hours, kWh) and the total amount of the bill (in $) for each billing cycle as well as the five-digit ZIP code of the premises. Only customers who were individually metered are included in the data set. For the purpose of this paper, we define a customer as a unique combination of premise and service account number. 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. While we have data covering additional years for two of the utilities, we limit the study to the years 2003 to 2006, to obtain equal coverage. Hereafter, we will refer to this data set as "billing data". Figure 3.1 displays the ZIP codes we have data for, which is the majority of the state.

Due to the difference in climate conditions across the state, California is divided into 16 building climate zones, each of which require different minimum efficiency building standards specified in an energy code.[2]. We expect this difference in build-

[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.

[2]Each climate zone has a representative city. These are for each of the climate zones: (1) Arcata, (2) Santa Rosa, (3) Oakland, (4) Sunnyvale, (5) Santa Maria, (6) Los Angeles, (7) San Diego, (8) El

Figure 3.1: Observed residential electricity consumption 2003 to 2006 and NOAA cooperative weather stations



Note: The map displays five-digit zip codes with available geographic boundaries.

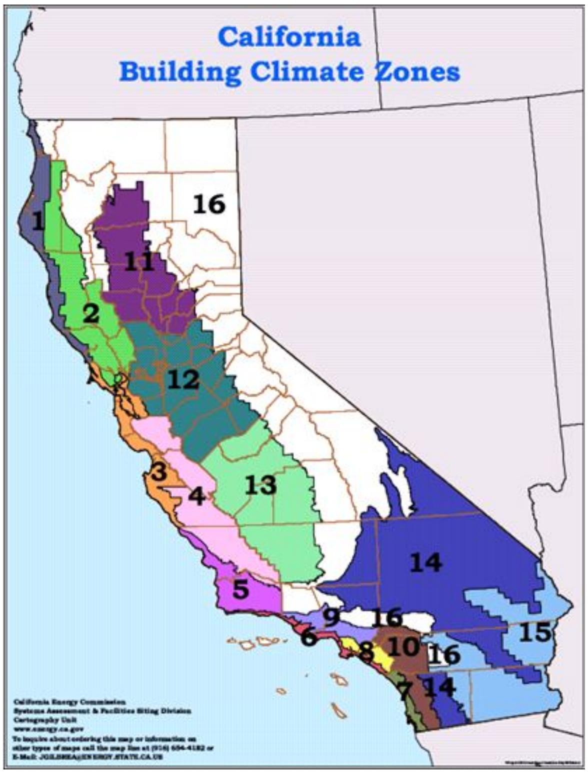
ing standards to lead to a different impact of temperature change on electricity consumption across climate zones. We will therefore estimate the impact of mean daily temperature on electricity consumption separately for each climate zone. We later empirically explore the sources of this variation in chapter 6. We assign each household to a climate zone via their five-digit ZIP code through a mapping, which we obtained from the California Energy Commission. The climate zones are depicted in Figure 3.2.

The billing data set contains 300 million observations, which exceeds our ability to conduct estimation using standard statistical software. We therefore resort to sampling from the population of residential households to conduct econometric estimation. We designed the following sampling strategy. First we only sample from households with regular billing cycles, namely 25-35 days in each billing cycle and which have at least 35 bills over the period of 2003-2006.[3] We also removed bills with

Toro, (9) Pasadena, (10) Riverside, (11) Red Bluff, (12) Sacramento, (13) Fresno, (14) China Lake, (15) El Centro, (16) Mount Shasta

[3]With the regular billing cycle, there should be 48 bills for the households in our sample during

Figure 3.2: California Energy Commission building climate zones



Source: California Energy Commission.

an average daily consumption less than 2 kWh or more than 80 kWh. The reason for this is our concern that these outliers are not residential homes, but rather vacation homes and small scale "home based manufacturing and agricultural facilities". Combined with the fact that our data does not contain single-metered multi-family homes, our sampling strategy is likely to result in a slight under representation of multifamily and smaller single family homes. These are more likely to be rental properties than larger single family units. Our results should be interpreted keeping this in mind.[4]

From the population subject to the restrictions above, we take a random sample from each ZIP code, making sure that the relative sample sizes reflect the relative sizes of the population by ZIP code. We draw the largest possible representative sample

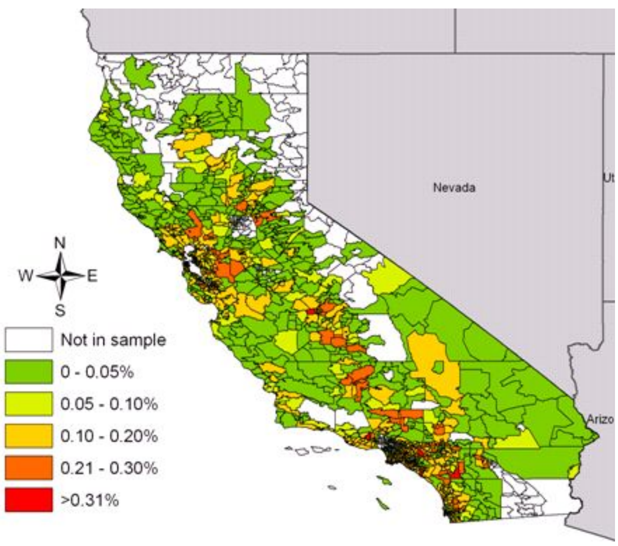
the period 2003 to 2006.

[4]After removing outlier bills, we compared the population average daily consumption of bills with billing cycles ranging from 25-35 days to the average daily consumption of bills for any length. The average daily consumption by climate zone in the subset of bills we sample from is roughly $\frac{1}{10}^{th}$ of a standard deviation higher than the mean daily consumption of the complete population including bills of any length.

from this population given our computational constraints. For each climate zone we test whether the mean daily consumption across bills for our sample is different from the population mean and fail to reject the null of equality, suggesting that our sampling is indeed random, subject to the sample restrictions discussed above. We proceed with estimation of our models by climate zone, which makes concerns about sampling weights mute. Figure 3.3 displays the spatial distribution of 2006 consumption shares across ZIP codes.

Finally, California has a popular program for low-income families - California Alternate Rates for Energy (CARE) - where program eligible customers receive a 20% discount on electric and natural gas bills. Eligibility requires that total household income is at or below 200% of federal poverty level. For the first set of models, we exclude these households from our sample. We then explore the robustness of our simulations by including these households in a separate simulation. The concern here is that omitting these smaller homes with lower HVAC saturation rates may lead to an overestimation of impacts.

Figure 3.3: Share of total residential electricity consumption for 2006 by five-digit zip code



No single ZIP code is responsible for more than 0.5% of total consumption. Table 3.1 displays the summary statistics of our consumption sample by climate zone. There is great variability in average usage across climate zones, with the central coast's (zone 3) average consumption per bill at roughly 60% that of the interior southern zone 15. The average electricity price is almost identical across zones, at 13 cents per kWh.

Table 3.1: Summary statistics for non-CARE households

		Usage per bill per billing cycle (Kwh)		Average price per billing cycle (Kwh)		Percentiles daily mean temperature distribution in sample (Degree Fahrenheit)				
No. of Obs	**No. of HH**	**Mean**	**S.D.**	**Mean**	**S.D.**	**1**	**5**	**50**	**95**	**99**
1,459,578	31,879	550	354	0.13	0.03	34.5	37.5	54.7	77.0	80.0
2,999,408	65,539	612	385	0.13	0.03	36.0	39.0	55.5	77.5	80.5
3,200,851	69,875	469	307	0.13	0.02	42.0	44.3	57.0	75.0	78.0
4,232,465	92,294	605	362	0.13	0.03	40.5	42.8	57.8	81.5	85.5
2,621,344	57,123	504	317	0.13	0.03	42.0	44.3	58.8	76.0	78.5
2,970,138	64,145	529	334	0.13	0.03	48.5	50.4	62.0	78.0	81.0
3,886,347	85,169	501	327	0.15	0.03	47.0	48.9	61.5	77.5	80.0
2,324,653	50,373	583	364	0.14	0.04	49.5	51.5	63.3	80.6	83.3
3,067,787	66,231	632	389	0.13	0.03	48.0	50.3	63.0	81.0	83.5
3,202,615	70,088	700	416	0.14	0.03	35.5	39.0	61.0	81.8	84.5
4,106,432	90,245	795	455	0.13	0.03	28.5	32.8	54.8	84.3	87.0
3,123,404	68,342	721	420	0.13	0.03	38.5	40.8	58.5	84.0	87.0
3,827,483	84,493	780	464	0.13	0.03	36.6	39.3	59.0	87.8	90.0
4,028,225	88,086	714	413	0.13	0.03	32.0	35.0	57.5	91.3	95.0
2,456,562	54,895	746	532	0.13	0.03	34.5	37.8	63.8	97.0	99.5
3,401,519	74,644	589	409	0.13	0.02	22.5	26.5	52.3	83.0	86.5

Notes: The table displays summary statistics for residential electricity consumption for the sample used in the estimation.

3.2 Weather Data

To generate daily weather observation to be matched with the household electricity consumption data, we use the Cooperative Station Dataset published by National Oceanic and Atmospheric Administration's (NOAA) National Climate Data Center (NCDC). The dataset contains daily observations from more than 20,000 cooperative weather stations in the United States, U.S. Caribbean Islands, U.S. Pacific Islands, and Puerto Rico. Data coverage varies by station. Since our electricity data cover the state of California for the years 2003-2006, the dataset contains 370 weather stations reporting daily data. In the dataset we observe daily minimum and maximum temperature as well as total daily precipitation and snowfall. Since the closest meaningful geographic identifier of our households is the five-digit postal ZIP code, we select stations as follows. First, we exclude any stations not reporting data in all years. Further we exclude stations reporting fewer than 300 observations in any single year and stations at elevations more than 7000 feet above sea level, which leaves us with 269 "valid" weather stations.[5] Figure 3.1 displays the distribution of these weather stations across the state. While there is good geographic coverage of weather stations for our sample, we do not have a unique weather station reporting data for each ZIP code. To assign a daily value for temperature and rainfall, we need to assign a weather station to each ZIP code. We calculate the Vincenty distance of a ZIP code's centroid to all valid weather stations and assign the closest weather station to that ZIP code. As a consequence of this procedure, each weather station on average provides data for approximately ten ZIP codes.

Since we do not observe daily electricity consumption by household, but rather monthly bills for billing periods of differing length, we require a complete set of daily weather observations. The NCDC data have a number of missing values, which we fill in using the following algorithm. First, we calculate the Vincenty distance of each ZIP code's geographic centroid to all qualifying weather stations. We then identify the ten closest weather stations to each centroid, provided that each is less than 50 miles from the monitor. Of these stations, we identify the "primary station" as the closest station reporting data for at least 200 days a year. We fill in missing values by first regressing, for observations in which the primary weather station was active, the relevant climate weather variable for the primary station onto the same variable for the remaining nine closest stations. We use the predicted values from that regression to replace missing values. Following this step, primary station observations are still missing whenever one of the remaining nine closest stations is also missing an observation. To estimate the remaining missing values, we repeat the above step with the 8 closest stations, then the 7 closest, etc. To check the performance of our algorithm, we conduct the following experiment. First, we select the set of data points

[5] The cutoff of 300 valid days is admittedly arbitrary. If we limit the set of weather stations to the ones providing a complete record, we would lose roughly half of all stations. We conducted robustness checks using different cutoff numbers and the results are robust.

for which the primary weather station has an observation. We then randomly set 10% of the temperature data for this station to missing. After applying the algorithm described above to this sample, we compare the predicted temperature data to the observations we had set aside. Even for observations in which a single additional weather station was used to predict a missing temperature, the correlation coefficient between actual and predicted temperatures exceeds 0.95. Plotting the actual and predicted series against each other provides an almost perfect fit. We therefore feel confident that our algorithm provides us with a close representation of the true data generating process for missing weather observations. We end up with a complete set of time series for minimum temperature, maximum temperature and precipitation for the 269 weather stations in our sample. For the remainder of our empirical analysis, we use these patched series as our observations of weather.[6]

There is an important caveat to using daily weather data when studying households' response to climate change. By using daily weather shocks we implicitly estimate individuals' response to changed *daily* temperatures. While climate change will affect daily temperatures on average, it is a more long run process and should be thought of as the long run moving average of weather. The estimated impacts for this reason may on the one hand be too high if individuals have lower cost options in the long run and relocate to cooler climates. The estimated impacts based on daily wether, on the other hand, may be too low if individuals adapt in the sense that areas which do not currently cool using electricity, start seeing a high degree of air conditioner penetration. The overall sign of the bias is not clear. Unfortunately it is not clear whether the perfect counterfactual to study this problem exists. One would require randomly assigned climate (not weather) to study this issue. This randomization would affect technology adoption. Electricity demand in turn is determined at the daily level by fluctuations in weather around a long run trend.

The second caveat is that it would be preferable to have a weather index, which counts all relevant dimensions of weather, such as minimum and maximum temperature, humidity, solar radiation and wind speed and direction. Unfortunately these indicators are not available for the vast majority of stations at the daily level. One could, however, estimate a response function using such an index for locations which have sufficient data. We leave this for future research.

3.3 Other Data

In addition to the quantity consumed and average bill amount, all we know about the households is the five-digit ZIP code in which they are 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

[6]We also tried an inverse distance weighting algorithm for filling in missing data and the results are almost identical.

(2006). The variables we will make use of are total population and average household income. The final sample used for estimation comprises households in ZIP codes which make up 81% of California's population. Table 3.2 displays summary statistics for all ZIP codes in California with registered residential population, broken down by whether we observe households in a given ZIP or not. We observe households for 1,325 ZIP codes and do not observe households for 239 ZIP codes. The 239 ZIP codes are not served by the three utilities, which provided us with access to their billing data. Table 3.3 shows that the ZIP codes in our sample are more populated, have larger households, are wealthier, and are at lower elevations. There seems to be no statistically significant difference in population, median age, or land area. Taking these differences into consideration is important when judging the external validity of our estimation and simulation results.

Finally, we will explore which observable characteristics of households are consistent with differences in the temperature repose function. We use the year 2000 long form census data for the state of California to calculate indicators of observable characteristics of the the average household or structure in that zip code. We obtain measures of the share of households using gas or electricity as heating fuel, year the average structure was built, the percent of urban households and the percent of rental properties.

Table 3.2: Summary statistics for zip codes in and out of sample

	Not in sample			In sample			
Variable	**No. of Obs.**	**Mean**	**S.D.**	**No. of Obs.**	**Mean**	**S.D.**	**Difference**
Population	239	19.83	20.86	1,325	20.39	20.67	0.56
Household size	239	2.66	0.6	1,325	2.79	0.60	0.14***
Household income	239	39.52	19.39	1,325	48.32	21.53	8.80***
House value	239	200.08	177.33	1,325	234.90	177.51	34.83***
Median age	239	36.92	7.34	1,325	3.85	7.50	-0.07
Elevation	239	1,081.45	1,526.95	1,325	439.63	737.94	-642***
Land area	239	69.66	130.12	1,325	68.05	140.45	-1.61

Chapter 4

Econometric Estimation

As discussed in the previous section, we observed each household's monthly electricity bill for the period 2003-2006. Equation (4.1) below shows our main estimating equation, which is a simple log-linear specification commonly employed in aggregate electricity demand and climate change impacts estimation (e.g., Deschênes and Greenstone 2007).

$$\log(q_{it}) = \sum_{p=1}^{k} \beta_p D_{pit} + \gamma Z_{it} + \alpha_i + \phi_m + \gamma_y + \varepsilon_{it} \tag{4.1}$$

$\log(q_{it})$ is the natural logarithm of household i's electricity consumed in kilowatt-hours during billing period t. 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 avoid the issue of having individuals moving to different structures with more or less efficient capital or residents with different preferences over electricity consumption moving in and out of a given structure. California's housing stock varies greatly across climate zones in its energy efficiency and installed energy consuming capital. We estimate equation (4.1) separately for each of the sixteen climate zones discussed in the data section, which are also displayed in Figure 3.1. The motivation for doing so is that we would expect the relationship between consumption and temperature to vary across these zones, as there is a stronger tendency to heat in the more northern and higher altitude zones and a stronger tendency to cool, but little heating taking place in the hotter interior zones of California.

The main variables of interest in this paper are those measuring temperature. The last five columns of Table 3.1 display the median, first, fifth, ninetieth, and ninety-fifth percentile of the mean daily temperature distribution by climate zone. The table shows the tremendous differences in this distribution across climate zones. The south eastern areas of the state for example, are significantly hotter on average, yet also have greater variances.

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 in weather (e.g., Schlenker and Roberts 2006). We achieve this by sorting each day's mean temperature experienced by household i into one of k temperature bins.[1] In order to define a set of temperature bins, there are two options found in the literature. The first is to sort each day into a bin defined by specific equidistant (e.g., 5 degree Fahrenheit) cutoffs. The second approach is to split each of the sixteen zones temperature distributions into a set of percentiles and use those as the bins used for sorting. The latter strategy allows for more precisely estimated coefficients, since there is guaranteed coverage in each bin. The equidistant bins strategy runs the risk of having very few observations in some bins and therefore leading to unstable coefficient estimation, especially at the extremes.

There is no clear guidance in the literature on which approach provides better estimates and we therefore conduct our simulations using both approaches. For the percentile strategy, 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 for each of the sixteen climate zones. The thresholds for each vary by climate zone. For the equidistant bins approach, we split the mean daily temperature for each household into a set of 5 degree bins. In order to avoid the problem of imprecise estimation at the tails due to insufficient data coverage, we require that each bin have at least 1% of the data values in it for the highest and lowest bin. The highest and lowest bins in each zone therefore contain a few values which exceed the 5 degree threshold.

For each household, bin definition and billing period we then counted the number of days the mean daily temperature falls into each bin and recorded this as D_{pit}. The main coefficients of interest to the later simulation exercise are the β_p's, which measure the impact of one more day with a mean temperature falling into bin p on the log of household electricity consumption. For small values, β_p's interpretation is approximately the percent change in household electricity consumption due to experiencing one additional day in that temperature bin.

Z_{it} is a vector of observable confounding variables which vary across billing periods and households. The first of two major confounders 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,

[1] We use mean daily temperature as our temperature measure. This allows a flexible functional form in a single variable. An alternate strategy we will explore in future work is separating the temperature variables into minimum and maximum temperature, which are highly correlated with our mean temperature measure.

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). The maximum likelihood approaches are computationally intensive and given our sample size cannot be feasibly implemented here. More importantly however, we do not observe other important characteristics of households (e.g., income) which would allow us to provide credible estimates of these elasticities. For later simulation we will rely on the income specific price elasticities provided by Reiss and White (2005), who used a smaller sample of more detailed data based on the national level RECS survey. We have run our models by including 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 bias 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.[2] The second major time varying confounder is precipitation in the form of rainfall. We calculate the amount of total rainfall for each of the 269 weather stations, filling in missing values using the same algorithm discussed in the previous section. We control for rainfall using a second order polynomial in all regressions.

The α_i are household fixed effects, which control for time invariant unobservables for each household. The ϕ_m are month-specific fixed effects, which control for unobservable shocks to electricity consumption common to all households. The γ_y are year fixed effects which control for yearly shocks common to all households. 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 $E[\varepsilon_{it} D_{pit} | D_{-pit}, Z_{it}, \alpha_i, \phi_m, \gamma_y] = 0$. Since we control for household fixed effects, identification comes from within household variation in daily temperature after controlling for shocks common to all households, rainfall, and average prices.

We estimate equation (4.1) for each climate zone using a least squares fitting criterion and a clustered variance covariance matrix clustered at the zip code.[3] Figure 5..1 and 5.2 plots the estimated temperature response coefficients for each of the climate zones against the midpoints of the bins for the percentile and equidistant bin approaches. The coefficient estimates are almost identical, which is reassuring. We do not display the confidence intervals around the estimated coefficients. The coefficients are so tightly estimated that for visual appearance, displaying the confidence intervals simply makes the lines appear thick. From this figure, several things stand out. First,

[2]The full set of estimation results are available upon request from the authors.

[3]Clustering along the time dimension would be desirable, but due to the temporal nesting structure of the billing dates not possible to our knowledge. We also used the White sandwich variance covariance matrix, which yielded smaller standard errors than the ones obtained from clustering by zip.

there is tremendous heterogeneity in the shape of the temperature response of electricity consumption across climate zones. Many zones have almost flat temperature response functions, such as southern coastal zones (5, 6, and 7). Other zones display a very slight negative slope at lower temperatures, especially the northern areas of the state (1, 2, and 11), indicating a decreased consumption for space heating as temperatures increase. California's households mostly use natural gas for space heating, which explains why for most areas we do not see a steeper negative slope at lower temperatures. This is consistent for a lower share of homes using electricity for heat in California (22%) than the national average (30%). Further, many of these electric heaters are likely located in areas with very low heating demand, given the high cost of using electricity for space heating compared to using natural gas. While there is use of electricity for heating directly, a significant share of the increased consumption at lower temperatures is likely to stem from the operation of fans for natural gas heaters. On the other end of the spectrum, for most zones in the interior and southern part of the state we note a significant increase in electricity consumption in the highest temperature bins (4, 8, 9, 10, 11, 12, 13, and 15). We further note that the relative magnitude of this approximate percent increase in household electricity consumption in the higher temperature bins varies greatly across zones as indicated by the differential in slopes at the higher temperatures across zones.

We now turn to simulating electricity consumption under different scenarios of climate change using these heterogeneous response functions as the underlying functional form relationship between household electricity consumption and temperature.

Chapter 5

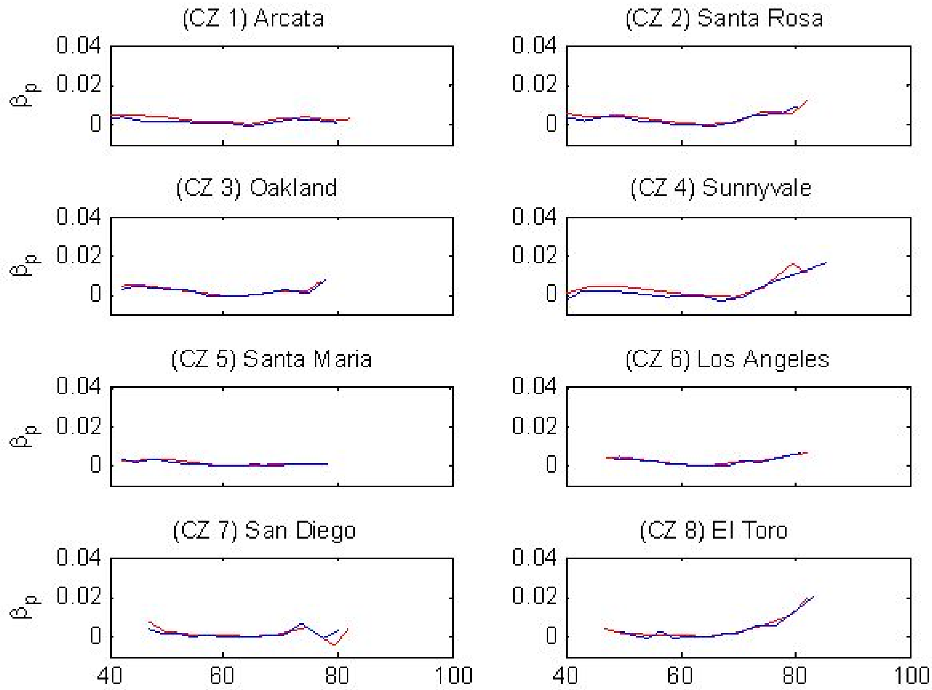
Simulations

In this section we simulate the impacts of climate change on electricity consumption under two different SRES emissions scenarios, three different electricity price scenarios, and three different population growth scenarios. We calculate a simulated trajectory of aggregate electricity consumption from the residential sector until the year 2100, which is standard in the climate change literature. To understand the impact of uncertainty surrounding these three different factors on aggregate consumption, we introduce them sequentially.

5.1 Temperature Simulations

To simulate the effect of a changing climate on residential electricity consumption, we require estimates of the climate sensitivity of residential electricity consumption as well as a counterfactual climate. In the simulation for this section we use the estimated climate response parameters shown in Figure 5.1 and 5.2. Using these estimates as the basis of our simulation has several strong implications. First, using the estimated β_p parameters implies that the climate responsiveness of consumption within climate zones remains constant throughout the century. This is a strong assumption, since we would expect that households in zones which currently do not require cooling equipment may potentially invest in such equipment if the climate becomes warmer. This would lead us to believe that the temperature responsiveness in higher temperature bins would increase over time. On the other hand, one could potentially foresee policy actions such as more stringent appliance standards, which improve the energy efficiency of appliances such as air conditioners. This would decrease the electricity per cooling unit required and shift the temperature response curve downwards in the higher buckets. We will deal with this issue explicitly in chapter 6.

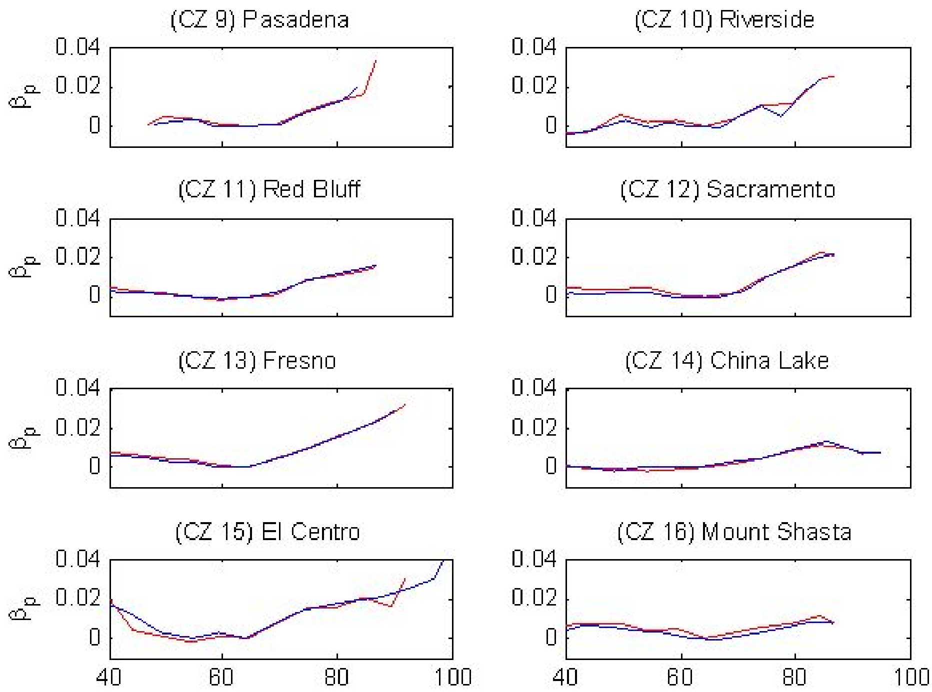
Figure 5.1: Estimated climate response functions for CEC climate zones 1 to 8.



Notes: The panels display the estimated temperature slope coefficients for each of the fourteen percentile bins (blue) and the equidistant bins (red) against the midpoint of each bin. The plots were normalized using the coefficient estimate for the 60 to 65 temperature bin. The title of each panel displays the name of a representative city for that climate zone.

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 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 "A2" and "B1" storylines in the 2000 Special Report on Emissions Scenarios (SRES) (IPCC 2000). The SRES study was conducted as part of the IPCC's Third Assessment Report, released in 2001. The A2 and B1 storylines and their quantitative representations represent two quite different possible trajectories for the world economy, society, and energy system, and imply divergent future anthropogenic emissions, with projected emissions in the A2 being substantially higher. The A2 scenario represents a "differentiated world", with respect to demographics, economic growth, resource use, energy systems, and cultural

Figure 5.2: Estimated climate response functions for CEC climate zones 9 to 16.



Notes: The panels display the estimated temperature slope coefficients for each of the fourteen percentile bins (blue) and the equidistant bins (red) against the midpoint of each bin. The plots were normalized using the coefficient estimate for the 60 to 65 temperature bin. The title of each panel displays the name of a representative city for that climate zone.

factors, resulting in continued growth in global CO_2 emissions, which reach nearly 30 gigatons of carbon (GtC) annually in the marker scenario by 2100. The B1 scenario can be characterized as a "global sustainability" scenario. Worldwide, environmental protection and quality and human development emerge as key priorities, and there is an increase in international cooperation to address them as well as convergence in other dimensions. A demographic transition results in global population peaking around mid-century and declining thereafter, reaching roughly 7 billion by 2100. Economic growth rates are higher than in A2, so that global economic output in 2100 is approximately one-third greater. In the B1 marker scenario, annual emissions reach about 12 GtC in 2040 and decline to about 4 GtC in 2100.

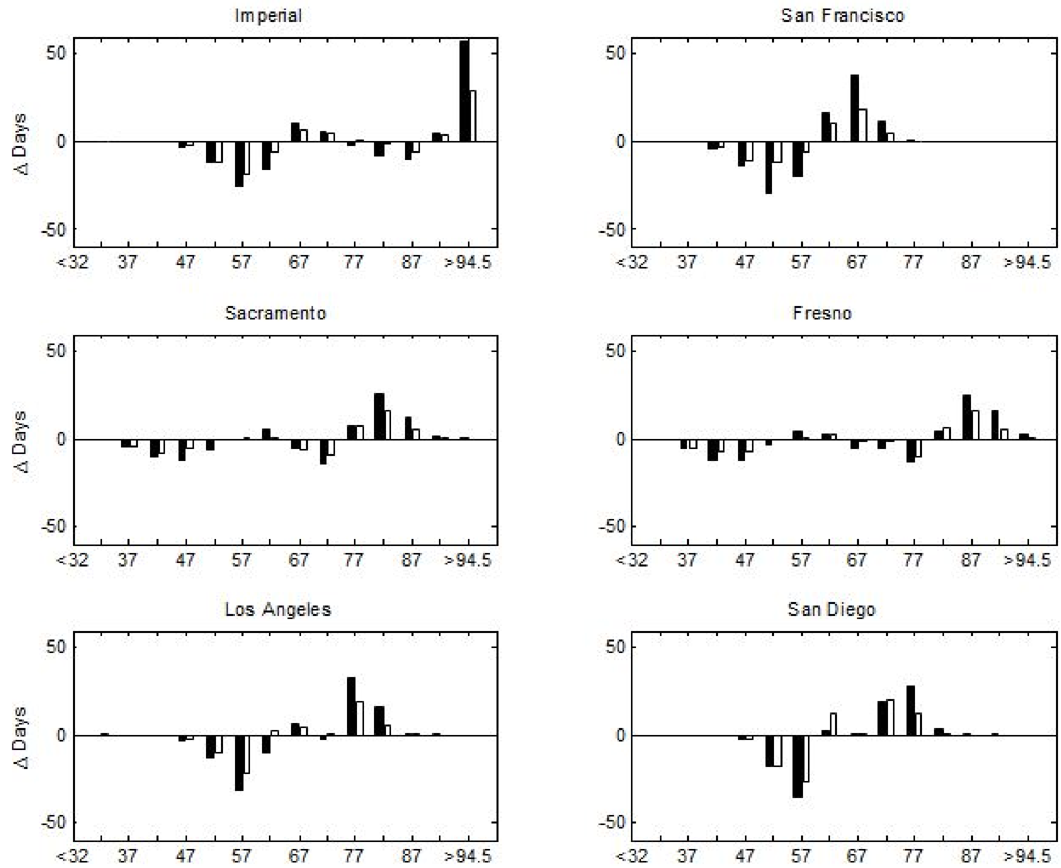
We simulate consumption for each scenario using the National Center for Atmospheric Research Parallel Climate Model 1 (NCAR). These models were provided to us in their downscaled version for California using the Bias Correction and Spatial Downscaling (BCSD) and the Constructed Analogues (CA) algorithms (Maurer and

Hidalgo 2008). There is no clear guidance in the literature as to which algorithm is preferable for impacts estimation. We therefore provide simulation results using both methods. 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:

$$\frac{q_{j,t+h}}{q_{j,t}} = \frac{\exp\left(\sum_{p=1}^{k} \hat{\beta}_{pj} D_{pj,t+h}\right)}{\exp\left(\sum_{p=1}^{k} \hat{\beta}_{pj} D_{pj,t}\right)} \tag{5.1}$$

We implicitly assume that the year fixed effect and remaining right hand side variables are the same for period $t+h$ and period t, which is a standard assumption made in the majority of the impacts literature. Figure 5.3 shows the change in the number of days spent in each 5 degree bin of the temperature distribution from 1980-1999 to 2080-2099 using the NCAR PCM forced by scenarios A2 and B1 for six selected California locations. A clear upward shift of the temperature distribution is apparent for all six locations. For locations with upward sloping temperature response functions, this entails increases in electricity consumption due to more days spent in higher temperature bins. Inspecting these graphs for all major urban centers in California, in addition to the six displayed here, confirms the pattern emerging from Figure 5.1 and 5.2. The areas with the steepest response functions at higher temperature bins happen to be the locations with highest increases in the number of high and extremely high temperature days. While this is not surprising, this correspondence leads to very large increases in electricity consumption in areas of the state experiencing the largest increases in temperature, which also happen to be the most temperature sensitive in consumption - essentially the southeastern parts of the state and the Central Valley.

Figure 5.3: Change in number of days in each 5-degree temperature bin for 2080 to 2099 relative to 1980 to 1999 for six selected California cities and IPCC SRES Scenario A2 (Black) and B1 (White) using the NCAR PCM with the constructed analogues downscaling method.



The first simulation of interest generates counterfactuals for the percent increase in residential electricity consumption by a representative household in each ZIP code. We feed each of the two climate model scenarios through equation (5.1) using the 1980-1999 average number of days in each temperature bin as the baseline. Figure 5.4 displays the predicted percent increase in per household consumption for the periods 2020-2039, 2040-2059, 2060-2079 and 2080-2099 using the NCAR PCM model forced by the A2 scenario using the percentile bins. Figure 5.5 displays the simulation results for the SRES forcing scenario B1.

Changes in per household consumption are driven by two factors: the shape of the weather-consumption relationship and the change in projected climate relative to the 1980-1999 period. The maps show that for most of California, electricity consumption at the household level will increase. The increases are largest for the Central Valley and areas in south eastern California, which have a very steep temperature response of consumption and large projected increases in extreme heat days. Simulation results for this model and scenario suggest that some ZIP codes in the Central Valley by the end of the century may see increases in household consumption in excess of 100%. The map also shows that a significant number of ZIP codes are expected to see drops in household level electricity consumption-even at the end of the current century. It is important to keep in mind that the current projections assume no change in the temperature electricity response curve. Specifically, the current simulation rules out an increased penetration of air conditioners in areas with currently low penetration rates (e.g., Santa Barbara) or improvements in the efficiency of these devices. The projected drops essentially arise from slightly reduced heating demand. We conduct a simulation below, which addresses this concern. Figure 5.5 displays the simulated household increase in electricity consumption by ZIP code for the lower emissions scenario B1. The maps display an almost identical spatial pattern, yet a smaller overall increase in consumption.

While changes in per household consumption are interesting, from a capacity planning perspective it is overall consumption that is of central interest from this simulation. We use the projected percent increase in household consumption by ZIP code and calculate the weighted overall average increase, using the number of households by ZIP code as weights, in order to arrive at an aggregate percent increase in consumption. The top panel of Table 5.1 displays these simulation results for aggregate consumption. Predicted aggregate consumption across all ZIP codes in our dataset ranges from an 18% increase in total consumption to 55% increase in total consumption by the end of the century. To put this into perspective, this represents an annual growth rate of aggregate electricity consumption between 0.17% and 0.44%, if all other factors are equal. These growth rates accelerate from period to period, as the number of extreme heat days predicted from the GCMs increases in a slightly non-linear fashion. For the first 20-year period, the simulated annual growth rates range from 0.10% per year to 0.29% per year. Since these simulations hold population constant, the correct comparison of these growth rates for the current simulation is therefore one with current growth in per capita household electricity consumption for California. Figure 5.6 depicts historical per capita electricity consumption since 1960 (EIA 2008). The average annual growth rate in per capita consumption during 1960-1973 was approximately 7% and slowed down to a remarkable 0.29% during 1974-1995. Growth rates during the last decade of available data have increased to a higher rate of 0.63%, and this difference in growth rates is statistically significant. The estimates from our simulation are lower than this growth rate and for the 2000-2019 period suggest that 26%-60% of this growth may be due to changing climate.

Figure 5.4: Simulated increase in household electricity consumption by zip code for the periods 2020 to 2039 (a), 2040 to 2059 (b), 2060 to 2079 (c), and 2080 to 2099 (d) in percent over 19801999 simulated consumption. Model NCAR PCM forced by IPCC SRES A2.

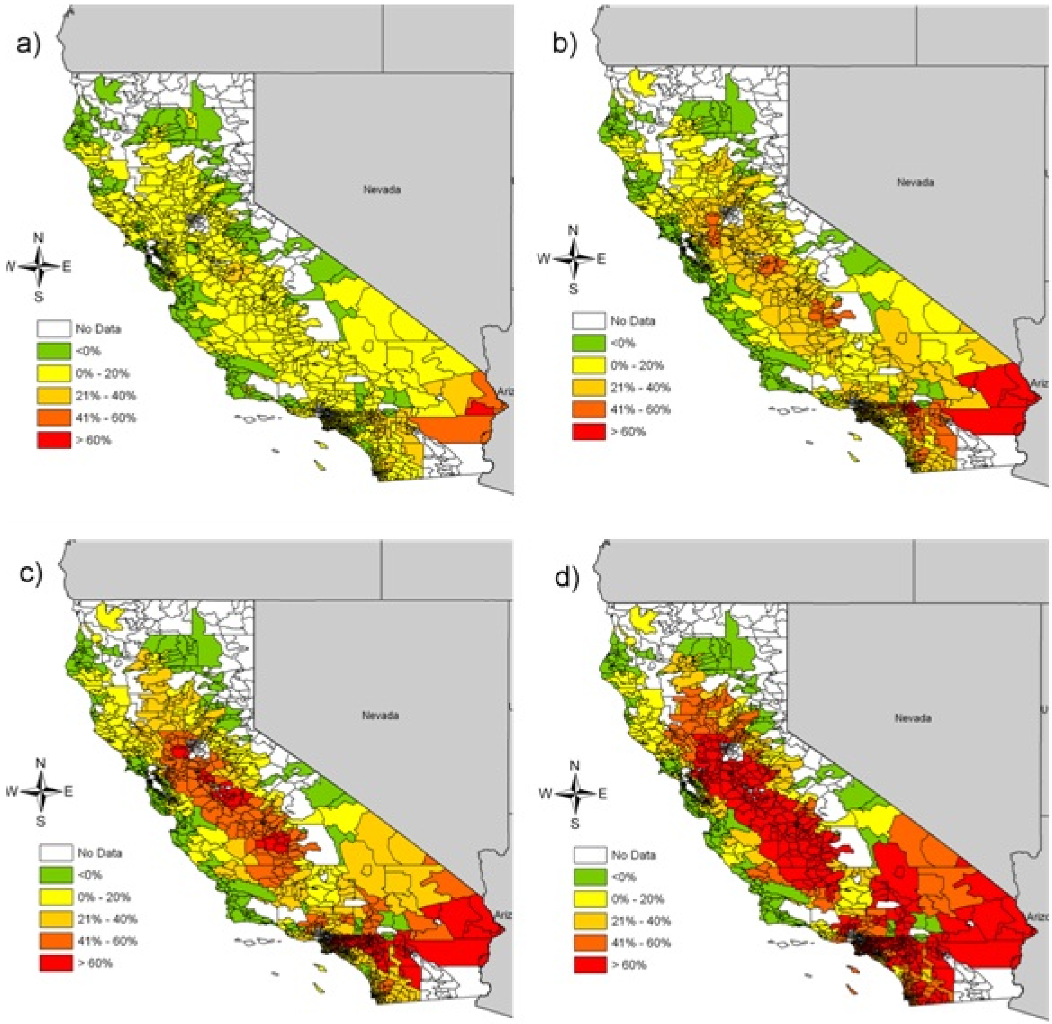


Figure 5.5: Simulated increase in household electricity consumption by zip code for the periods 2020 to 2039 (a), 2040 to 2059 (b), 2060 to 2079 (c), and 2080 to 2099 (d) in percent over 19801999 simulated consumption. Model NCAR PCM forced by IPCC SRES B1.

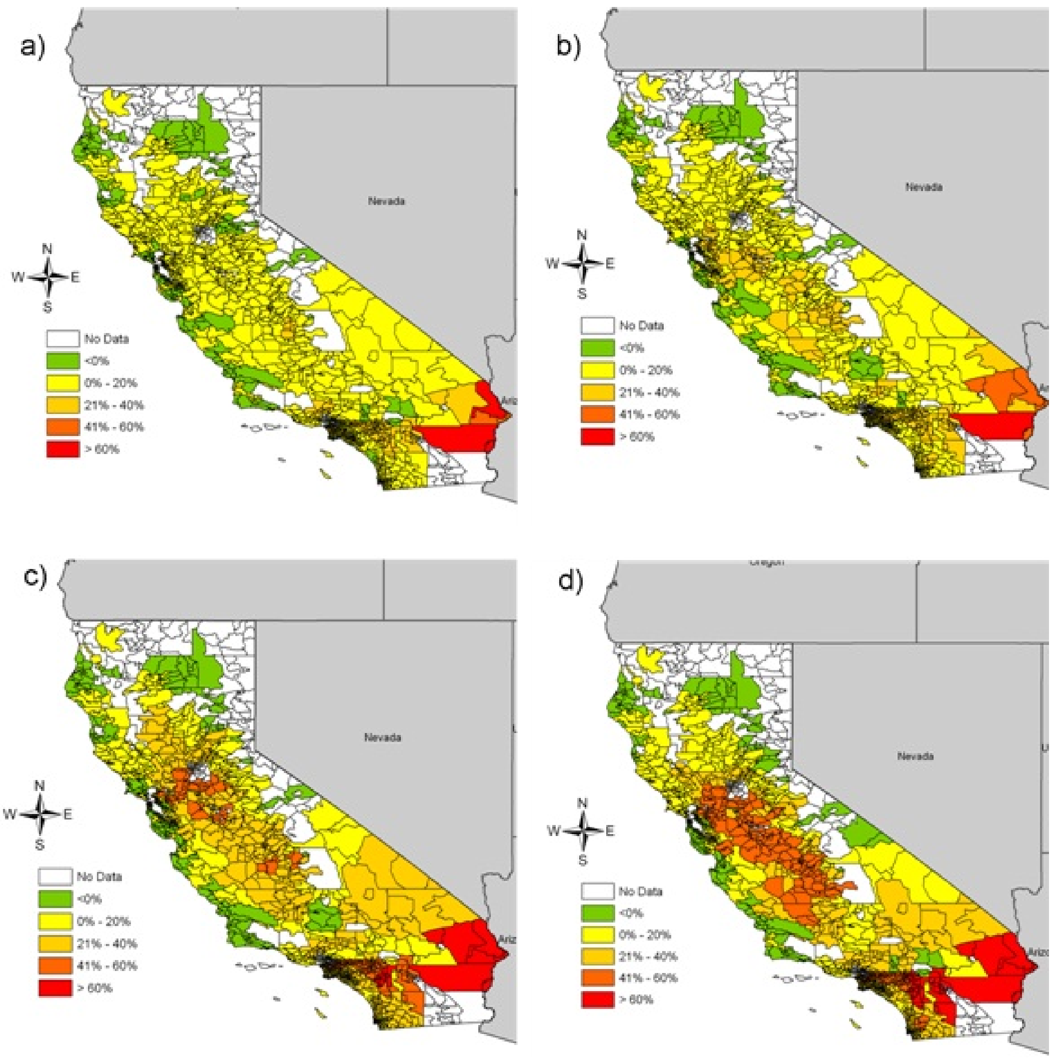
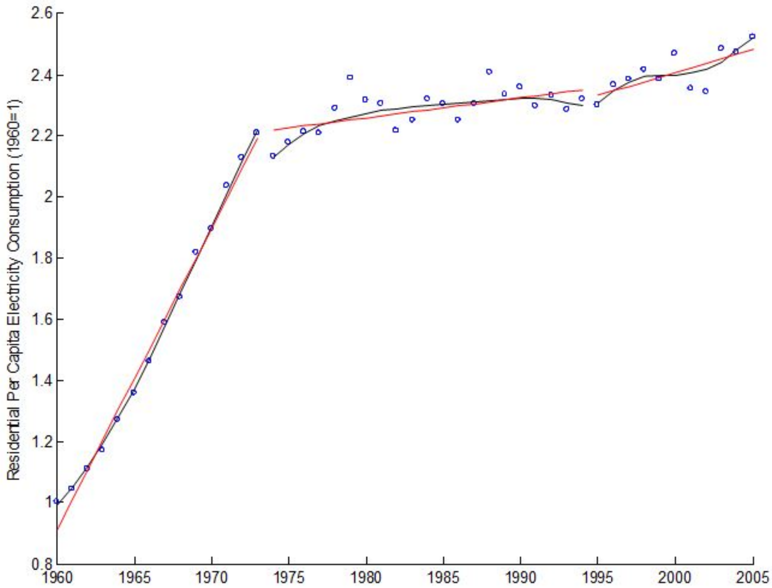


Table 5.1: Simulated percent increase in residential electricity consumption relative to 1980 to 2000 for the constant, low price, and high price scenarios

		Equidistant-bins				Percentile-bins			
	Down-scaling	**BCSD**		**CA**		**BCSD**		**CA**	
	IPCC Scenario	**A2**	**B1**	**A2**	**B1**	**A2**	**B1**	**A2**	**B1**
	Price Increase								
2000-19	**0%**	5%	2%	5%	3%	6%	3%	5%	3%
2020-39	**0%**	5%	8%	7%	8%	6%	9%	7%	8%
2040-59	**0%**	15%	9%	17%	10%	17%	11%	17%	10%
2060-79	**0%**	24%	15%	28%	16%	28%	17%	28%	16%
2080-99	**0%**	48%	18%	50%	20%	55%	21%	50%	20%
2000-19	**0%**	5%	2%	5%	3%	6%	3%	5%	3%
2020-39	**30%**	-6%	-3%	-5%	-4%	-5%	-3%	-4%	-3%
2040-59	**30%**	3%	-2%	3%	-2%	6%	-1%	5%	-1%
2060-79	**30%**	11%	3%	11%	2%	15%	5%	15%	4%
2080-99	**30%**	33%	6%	29%	4%	39%	9%	35%	7%
2000-19	**0%**	5%	2%	5%	3%	6%	3%	5%	3%
2020-39	**30%**	-6%	-3%	-5%	-4%	-5%	-3%	-4%	-3%
2040-59	**60%**	-9%	-13%	-8%	-13%	-6%	-12%	-7%	-12%
2060-79	**60%**	-1%	-9%	-1%	-10%	2%	-7%	2%	-8%
2080-99	**60%**	18%	-6%	15%	-7%	24%	-4%	20%	-5%

Figure 5.6: California residential per capita electricity consumption.



Source: Authors calculations based on EIA (2008) SEDS data.

5.1.1 Temperature and Price Simulations

The assumed flat prices from the previous section should be considered as a comparison benchmark. It is meaningful and informative to imagine climate change imposed on today's conditions. It is worth pointing out, however, that real residential electricity prices in California have been on average flat since the early-mid 1970s spike. In this section we will relax the assumption of constant prices and provide simulation results for increasing electricity prices under a changing climate.

While we have no guidance on what will happen to retail electricity prices 20 years or further out into the future, we construct two scenarios. The first scenario we consider is a discrete 30% increase in real prices starting in 2020 and remaining at that level for the remainder of the century. This scenario is based upon current estimates of the average statewide electricity rate impact by 2020 of AB 32 compliance combined with natural gas prices to generators within the electric power sector. These estimates are based on analysis commissioned by the California Public Utilities Commission. This scenario represents the minimum to which California is committed in the realm of electricity rates. This scenario could be interpreted as one assuming very optimistic technological developments post 2030, implying that radical CO_2 reduction does not entail any cost increases, or as a California and worldwide failure to pursue dramatic

CO_2 reductions such that California's AB 32 effort is not expanded. The second scenario we consider is one where electricity prices increase by 30% in 2020 and again by 30% in 2040 and remain at that level thereafter. We consider the additional increase in mid-century price in essence as an "increasing marginal cost" story. Under this scenario, AB 32 is successfully implemented and a path towards achieving the 2050 targets is put in place. These additional steps are assumed to be proportionally more expensive.

To simulate the effects of price changes on electricity consumption, we require good estimates of the price elasticity of demand. In this paper we rely on the estimates of mean price elasticity provided by Reiss and White (2005). Specifically, they provide a set of average price elasticities for different income groups, which we adopt here. Since we do not observe household income, we assign a value of price elasticity to each ZIP code based on the average household income for that ZIP code. Households are separated into four buckets, delineated by \$18,000, \$37,000, \$60,000 with estimated price elasticities of -0.49, -0.34, -0.37, and -0.29 respectively. It is important to note that these price elasticities are short-run price elasticities. These are valid if one assumes a sudden increase in prices, as we do in this paper. To our knowledge, reliable long-term price elasticities based on micro data for California are not available, but in theory they are larger than the elasticities used in this paper.

The second panel in Table 5.1 presents the simulation results under the two different scenarios of climate change given a sudden persistent increase in electricity prices in the year 2020. Given the range of price elasticity estimates, it is not surprising that the simulated increases in residential electricity consumption for the first period after the price increase are roughly 6%-12% lower than the predicted increases given constant prices. For the NCAR model under both considered forcing scenarios and both downscaling algorithms, the path of electricity consumption under these price scenarios returns to levels below its 1980-2000 mean for the 2020-2040 period, given this assumed price trajectory.

The third panel in Table 5.1 presents the simulation results for both forcing scenarios and downscaling methods given the high price scenario. Given the significant increase in prices after 2020 and again in 2040, the consumption trajectory stays flat for the entire simulation period using the NCAR model for the B1 scenario. The higher forcing scenario A2 shows a relatively flat trajectory, yet still predicts significant increases in consumption for the last decades of the century-even in the face of these higher prices. It is important to note that these effects are conditional on the estimated price elasticities being correct. Smaller elasticities would translate into price based policies, such as taxes or cap and trade systems, being less effective at curbing demand compared to standards.

5.1.2 Temperature and Population

California has experienced an almost seven-fold increase in its population since 1929 (BEA 2008). California's population growth rate over that period (2.45%) was more than twice that of the national average (1.17%). 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 fifty-five years out of sample, 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.

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% 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% annual growth rate and 2.3-fold increase in total population. The low series is equivalent to a 0.18% growth rate and results in a population 18% 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 5.2 displays the simulated aggregate electricity consumption given the three population growth scenarios. This table holds prices constant at the current level and therefore presents a "worst case scenario". It is not surprising to see that population uncertainty has much larger consequences for simulated total electricity consumption compared to uncertainty over climate or uncertainty over prices. The simulations for the low forcing scenario B1 and the low population growth scenario show 65%-70% increase in residential electricity consumption. The same figure for the medium growth scenario predicts a 179%-189% increase in consumption and the worst case scenario predicts a 350% increase in consumption. If we consider the A2 forcing, the predicted low population average increase in consumption is a 118% increase or a 478% increase for the high population growth scenario.

Table 5.2: Simulated percent increase in residential electricity consumption relative to 1980 to 2000 for the low, middle, and high population scenarios

	Equidistant-bins				Percentile-bins			
Down-scaling	**BCSD**		**CA**		**BCSD**		**CA**	
IPCC Scenario	**A2**	**B1**	**A2**	**B1**	**A2**	**B1**	**A2**	**B1**
Low population growth scenario								
2000-19	17%	13%	16%	14%	18%	14%	16%	15%
2020-39	31%	34%	33%	34%	32%	35%	34%	35%
2040-59	48%	41%	50%	41%	52%	42%	53%	42%
2060-79	66%	52%	68%	51%	72%	55%	73%	54%
2080-99	113%	65%	113%	65%	124%	70%	123%	70%
Medium population growth scenario								
2000-19	19%	15%	18%	16%	19%	16%	18%	16%
2020-39	48%	52%	51%	52%	50%	54%	52%	53%
2040-59	99%	88%	101%	89%	104%	91%	105%	91%
2060-79	154%	133%	157%	133%	164%	139%	166%	138%
2080-99	258%	179%	257%	179%	277%	189%	275%	188%
High population growth scenario								
2000-19	23%	19%	22%	20%	23%	20%	22%	20%
2020-39	64%	68%	66%	66%	66%	70%	68%	69%
2040-59	135%	123%	137%	123%	141%	126%	142%	125%
2060-79	240%	212%	243%	211%	252%	219%	254%	218%
2080-99	464%	342%	462%	342%	495%	357%	490%	356%

5.1.3 CARE Customers

All of the results presented in the paper so far have excluded CARE customers from the estimation sample. One potential concern is that these households live on fewer square feet, are more likely to be renting, have lower average use and lower HVAC saturation rates. This would suggest that the temperature response for these households is potentially lower than for the households in the full sample. The number of CARE households in California is large. SCE reports over 1 million customers on CARE, which is roughly one quarter of residential accounts. For PG&E and SDG&E the share of accounts is roughly 20%. We therefore separately sample from only the CARE households by ZIP code, adopting the same sampling restrictions as in the non-CARE sample. We then estimate temperature response functions by climate zone, which are slightly less steep in the higher temperature bins. We then conduct the simulations for the CARE households separately. To obtain an estimate of the overall impacts, when we include CARE, we weight impacts for each ZIP code by the share of CARE to non-CARE households in that ZIP code. Table 5.3 reports these results for the BCSD downscaling algorithm and equidistant bin simulations. As suspected, the CARE households are slightly less affected by higher temperatures, yet the overall weighted average is very close to the simulations presented in Table 5.1.

Table 5.3: Simulated percent increase in residential electricity consumption relative to 1980 to 2000 for CARE and non-CARE households.

Down-scaling IPCC Scenario		NON-CARE Equidistant-bins BCSD A2	NON-CARE Equidistant-bins BCSD B1	CARE Equidistant-bins BCSD A2	CARE Equidistant-bins BCSD B1	WEIGHTED Equidistant-bins BCSD A2	WEIGHTED Equidistant-bins BCSD B1
	Price Increase						
2000-19	**0%**	5%	2%	4%	2%	5%	2%
2020-39	**0%**	5%	8%	4%	6%	5%	7%
2040-59	**0%**	15%	9%	12%	8%	14%	9%
2060-79	**0%**	24%	15%	20%	12%	23%	14%
2080-99	**0%**	48%	18%	39%	15%	46%	17%
2000-19	**0%**	5%	2%	4%	2%	5%	2%
2020-39	**30%**	-6%	-3%	-6%	-4%	-6%	-4%
2040-59	**30%**	3%	-2%	1%	-3%	2%	-2%
2060-79	**30%**	11%	3%	8%	1%	10%	2%
2080-99	**30%**	33%	6%	25%	3%	31%	5%
2000-19	**0%**	5%	2%	4%	2%	5%	2%
2020-39	**30%**	-6%	-3%	-6%	-4%	-6%	-4%
2040-59	**60%**	-9%	-13%	-11%	-14%	-9%	-13%
2060-79	**60%**	-1%	-9%	-5%	-10%	-2%	-9%
2080-99	**60%**	18%	-6%	-11%	-9%	16%	-7%

Chapter 6

Adaptation Simulations

As mentioned at the beginning of this chapter, all previous simulation exercises we have conducted so far assumed that the temperature response functions are fixed for each climate zone until the end of the century. This implicitly assumes that people do not adapt to a changing climate, which is a crucial and potentially non-credible assumption. If the coastal areas of California will experience higher mean temperatures and more frequent extreme heat events, it is likely that newly constructed homes will have built in central air conditioning. Further, owners of existing homes may install air conditioning equipment ex-post. This type of adaptation would result in a stronger temperature response at higher temperatures, whereby the temperature elasticity in the highest bins would increase over time. On the other hand, forward-looking planners and policy makers may put in place more stringent building codes for new construction combined with more stringent appliance standards, which would decrease the energy intensity of existing capital and homes. California has a long history of these policies and is considered a worldwide leader in these energy efficiency policies. These more stringent policies are designed to offset future increases in consumption. Bottom up engineering models by design can capture the impact of building- and device-specific changes due to regulations on energy consumption. Their drawback, however, is that they have to rely on a large number of assumptions regarding the composition of the housing stock and appliances, as well as making behavioral assumptions about the individuals using them.

While we cannot conduct a detailed simulation incorporating specific policy changes, we conduct the following thought experiment. Our baseline simulation has assumed that each climate zone maintains its specific response function throughout the remainder of the century. To bound how important the heterogeneity in the response function is to the aggregate simulation results, we design an “almost best case” scenario, where we assume that all zones have the response function of coastal San Diego (Zone 7). This zone’s response function is relatively flat. The left panel of Table 6.1 shows the simulation results assuming this optimistic scenario.

Table 6.1: Simulated percent increase in residential electricity consumption relative to 1980 to 2000 assuming a common low (Zone 7) and high (Zone 12) temperature response function

	Zone 7		**Zone 12**	
	A2	**B1**	**A2**	**B1**
2000-19	1%	-1%	13%	7%
2020-39	1%	1%	5%	7%
2040-59	2%	1%	29%	13%
2060-79	2%	1%	57%	28%
2080-99	3%	0%	122%	40%

Worst-case increases under forcing A2 results in a 3% increase in electricity consumption by the end of the century, which is essentially flat compared to the baseline simulation shown in Table 5.1. Next we come up with an "almost worst case" scenario, where we let all of California adopt the response function of Zone 12, the Central Valley. The right panel of Table 6.1 shows the results from this simulation. The overall increases in simulated electricity consumption are more than twice those of the baseline scenario across simulations considered in Table 5.1. The large impact of the assumed temperature responsiveness function on overall simulated residential electricity consumption underlines the importance of improving energy efficiency of buildings and appliances.

6.1 Heterogeneity in temperature response functions

The major finding in the paper so far is the heterogeneity in temperature response of residential electricity consumption across climate zones. While geographic location clearly plays an important role in determining this responsiveness, we wish to study whether there are household or structure characteristics, which help explain some of this difference in temperature response. We therefore construct a state-wide sample by sampling 10% of the households from each of the sixteen climate zone specific datasets used above. We restrict ourselves to non-Care customers in this exercise. We construct 10 percentile temperature bins, where the cutoffs are at every 10^{th} percentile of the California wide temperature distribution for the years 2003-2006. The smaller number of bins and percentile approach guarantee that there are enough observations in the extreme bins at meaningful cutoff points.

We then slice the dataset above along several dimensions in order to see whether the temperature response varies with certain variables of interest from the census 2000 Summary File 3 (SF 3). Specifically, for each indicator, we divide this sample into 2 groups, a "low group" and a "high group", based on the value of the variables of interest. The variables of interest and percentiles used in estimation are:

1. Percentage of household using electricity as heating fuel
   - Low group: households in zip code with this variable <= 30%
   - High group: households in zip code with this variable >= 60%
2. Percentage of household using gas as heating fuel
   - Low group :households in zip code with this variable <= 40%
   - High group: households in zip code with this variable >= 60%
3. Percentage of households in an urban area
   - Low group :households in zip code with this variable <= 40%
   - High group: households in zip code with this variable >= 60%
4. Median year of structure built (base year is 1939 =1)
   - Low group (older building) : ZIP codes with median age home < 20 years
   - High group (newer building): ZIP codes with median age home > 40 years

For each variable of interest, we estimate the same models as previously, while making sure that we are making a fair comparison across groups. For our regressions we therefore limit the sample for both groups to those households with median household income between 40-60% of the distribution of census 2000 zip-code-level median household income.

For each variable of interest, we plot the estimated coefficients for each temperature bin against their mid-bin temperature. Each of the graphs has 2 sets of lines, one for "low group" (red) and the other one for "high group" (black). We also plot the 95% confidence intervals for each group. Figure 6.1 plots the response functions for households in zip codes with a high penetration of electricity as the major heating fuel against the response functions for households from zip codes with a low penetration of electricity of a heating fuel. The difference is drastic and statistically significant. The ZIP codes using electricity as the major source of heat have significantly higher electricity consumption at low temperatures, while the low penetration zip codes have an almost flat response. The panel below displays the figure for natural gas. It is switched, which is not surprising, given that electricity and natural gas are the two major heating fuels in California. In the top panel it is also noteworthy that households with higher electric heating have a drastically higher temperature response at high temperatures.

Figure 6.1 displays the temperature response functions for older houses versus newer houses in the top panel. At the low temperature spectrum, newer houses seem to require less electricity to heat compared to older houses, which is consistent with the belief that building codes provide better insulation. At the high end of the temperature spectrum, older and newer houses appear to have an almost identical temperature response, which may have to do with the higher penetration of central air conditioning in the newer and larger homes. The bottom panel of Figure 6.1 displays the temperature response for houses located in mostly urban ZIP codes versus the temperature response of households located in mostly rural ZIP codes. The difference is quite drastic, with rural households having an almost flat temperature response function and urban households having the typical U-shaped response. This finding is due to the fact that much of the central valley and the greater LA area are considered urban.

Figure 6.1: Temperature response for households by major heating fuel.

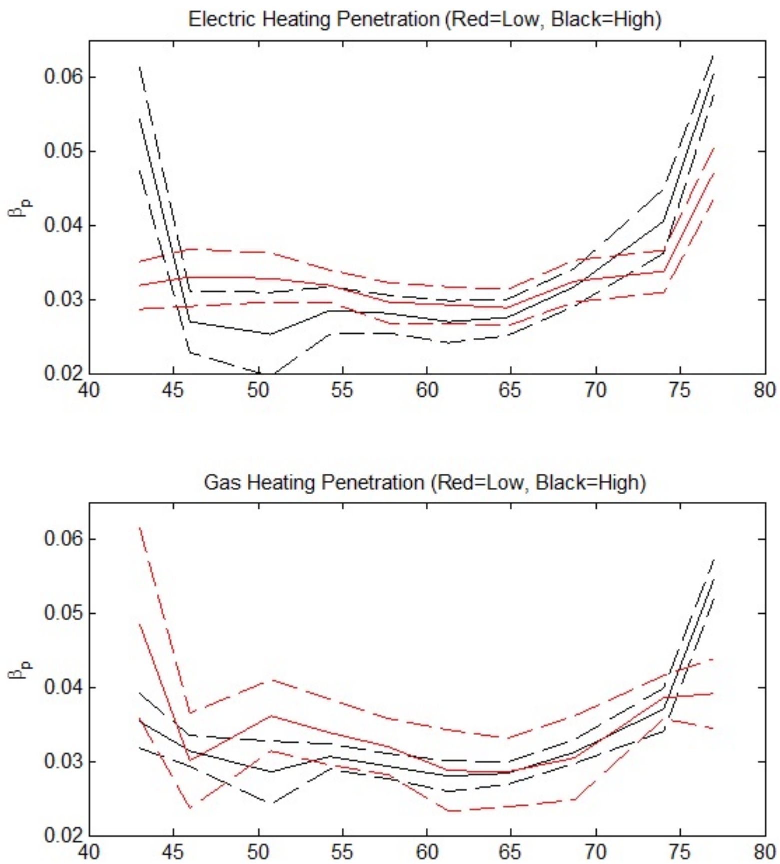
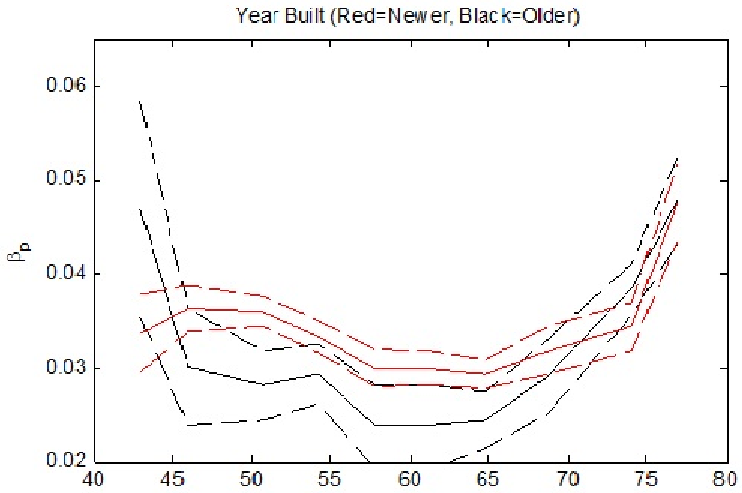
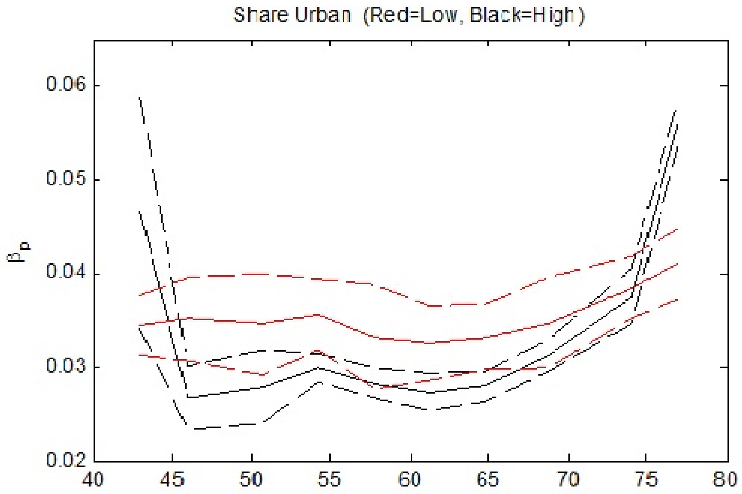


Figure 6.2: Temperature response for households by year built and urban location.





Chapter 7

Conclusions

This study has provided the first estimates of California's residential electricity consumption under climate change based on a large set of panel micro-data. We use random and therefore exogenous weather shocks to identify the effect of weather on household electricity consumption. We link climate zone specific weather response functions to a state of the art downscaled global circulation model to simulate growth in aggregate electricity consumption. We further incorporate potentially higher prices and population levels to provide estimates of the relative sensitivity of aggregate consumption to changes in these factors. Finally we show estimates of aggregate consumption under an optimistic and pessimistic scenario of temperature response.

There are several novel findings from this paper. First, simulation results suggest much larger effects of climate change on electricity consumption than previous studies. This is largely due to the highly non-linear response of consumption at higher temperatures. Our results are consistent with the findings by Greenstone and Deschênes (2007). They find a slightly smaller effect using national data. It is not surprising that impacts for California, a state with a smaller heating demand (electric or otherwise), would be bigger. Second, temperature response varies greatly across the climate zones in California - from flat to U-shaped to hockey stick shaped. This suggests that aggregating data over the entire state may ignore important nonlinearities, which combined with heterogeneous climate changes across the state may lead to underestimates of future electricity consumption. Third, population uncertainty leads to larger uncertainty over consumption than uncertainty over climate change. Finally, policies aimed at reducing the weather sensitivity of consumption can play a large role in reducing future electricity consumption. Specifically, region specific HVAC standards may play a significant role in offsetting some of the projecting increases in consumption.

Part II

The Impacts of State Level Building Codes on Residential Electricity Consumption

Chapter 8

Introduction

U.S. residential electricity consumption, which accounted for 37% of total electricity consumption in 2006, has increased by 570% or on average by 4.2% annually from 1960 to 2007. In 2007, the sector contributed about 21 percent of U.S. CO_2 emissions from fossil fuel combustion, more than two-thirds of which are due to electricity consumption (EIA, 2008).

In order to meet ever increasing demand, utilities have added generating capacity, while at the same time implementing measures to slow the growth of energy demand. One set of such measures is in the form of energy efficiency policies which are thought to reduce the demand for electricity and carbon emissions. These policies can be categorized into four main types: energy efficiency standards (*e.g.* building energy codes and appliance standards), financial incentives for energy efficient investment (*e.g.* rebate programs), information and voluntary programs (*e.g.* advanced metering), and management of government energy use. The major policies that directly affect residential electricity consumption are appliance standards and building energy codes. Appliance standards require manufacturers to meet minimum energy efficient standards to sell their product in the geographic area of adoption (e.g. state). Building energy codes require newly constructed buildings as well as modified existing buildings to meet certain engineering specifications relevant to energy consumption.

In the residential sector, demand for electricity is derived from the use of electrical appliances which provide energy services such as refrigeration, heating and cooling. According to a 2001 household energy consumption survey, appliances (*e.g.* air conditioners, refrigerators, space and water heating systems and lights) are the largest users of electricity in the average U.S. household, consuming approximately two-thirds of all the residential electricity (EIA, 2001). As such, the energy efficiency of these appliances, defined as the units of energy per unit of service provided, is a major factor determining household and aggregate electricity consumption.

Residential building energy codes provide minimum building requirements for heating and cooling systems and for the housing envelope that lead to energy savings. For example, with careful building envelope design, good insulation and window glazing selection, builders can significantly downsize or even eliminate heating and cooling equipment or reduce the frequency and/or intensity of its use. More energy efficient buildings and appliances are designed to offset some of the otherwise predicted demand for energy.

The effectiveness of these codes and standards has been widely studied across disciplines. The vast majority of studies on energy savings due to these codes and standards are ex ante studies conducted by engineers. These studies have the advantage of being able to simulate changes in derived demand from specific policy induced scenarios of technological change at the building or appliance level. In order to simulate future consumption patterns one has to make detailed assumptions about the adoption of each technology and its usage, which are two factors not well understood empirically. A few recent studies attempt to overcome these issues by decomposing aggregate demand for a given state (e.g. Sudarshan and Sweeney, 2008) or across states (e.g. Horowitz, 2008) into price, income, climate and policy effects. Both studies show significant effects of state level policies on electricity consumption.

Our current study makes three specific contributions to this literature. First, instead of focusing on broadly defined energy efficiency policies, we quantify the effect of a specific and widely applied policy tool - residential building codes - on per capita residential electricity consumption. Second, we econometrically identify the effect of building codes on residential electricity consumption by exploiting temporal and spatial variation in the introduction of state level building codes and new construction instead of adopting a bottom-up modeling approach. The econometric approach has the advantage that it uses observed consumption data ex post, which embeds the behavioral response of the consumer. Finally, we control for the endogeneity of price and our policy variable by using an instrumental variables estimation strategy. Our findings should be of broader interest, since national residential and commercial building codes are the core energy efficiency policies in Waxman-Markey climate bill, which the house passed on June 26, 2009.

The next chapter briefly discusses the history of building codes and standards. Chapter 11 presents empirical model and describes the data. Chapter 12 presents estimation results. Chapter 13 concludes.

Chapter 9

Background

The energy crisis and growing environmental concerns of the late 1960s and 1970s were key factors stimulating the development of public policies aimed at promoting energy efficiency and conservation in the United States, primarily through technology regulation. California's Warren-Alquist Act, enacted in 1974, established the California Energy Commission and granted it authority to introduce and enforce environmental criteria in the production and consumption of energy. Energy efficiency standards for residential and non-residential buildings were enacted in 1978 through Title 24 of the California Code of Regulations. At the federal level, the 1975 Energy Policy and Conservation Act was amended in 1978 to include, as a condition of receiving federal funding, requirements for state conservation and efficiency programs including building energy codes. Through the 1980s, a number of states adopted codes based on the ASHRAE (American Society of Heating, Refrigerating, and Air Conditioning Engineers) code 90-1075. Other states adopted the Model Energy Code (MEC) developed by the Council of American Building Officials (CABO) (Howard and Prindle 1991). In 1992, the enactment of the Energy Policy Act included a provision for states to review and/or revise their residential building codes regarding energy efficiency to meet the CABO Model Energy Code. The MEC has since been revised and updated, and its successor is the International Energy Conservation Code (IECC). Currently, all states except Hawaii, Kansas, Mississippi, Missouri, Illinois and South Dakota have implemented a statewide version of building codes. Title II of The American Clean Energy and Security Act of 2009, passed by the U. S. House of Representatives in June, includes the establishment of national energy codes for residential and commercial buildings, with the residential codes based upon the IECC 2006 code.

There is a large literature on the underlying policy rationale and the economic logic of technology-based energy efficiency regulations, including building codes. It is primarily addressed to the issue of the so-called "energy efficiency gap", the difference between observed efficiency investments and those deemed cost-effective by engineering (life-cycle cost) criteria. (Sanstad et al. 2006 is a recent overview.) Estimation of this difference is also the basis for prospective "efficiency potential" analyses (National Action Plan for Energy Efficiency 2007). By contrast, the empirical literature on ex post estimation of electricity and natural gas consumption reductions resulting from energy efficiency policies is relatively limited. Gillingham et al. (2006) estimate U. S. cumulative savings from efficiency policies and programs, but exclusive of building energy codes.

Empirical ex post estimation of the energy savings from building codes and standards poses several challenges, whether at the individual building level or at higher levels of aggregation. For example, California has established regulatory requirements and guidelines for ex post measurement of energy savings from utility demand-side management programs, which use information gathered independently of the ex ante engineering savings estimates used to design and implement the programs (Sanstad, 2007). By contrast, such measurements for buildings have only recently begun to emerge in a research context, as technology and methods have improved. In turn, even as such data become more commonly available, aggregate, long-term retrospective savings estimates based upon building-level information must rely on construction of a counter-factual - that is, a "version of history" without the policies - for comparison. This requires considerably more than building code data alone. Thus, for example, the California Energy Commission's well-known estimates of historical savings due to state efficiency policies and programs are based upon an energy demand simulation model, which is run over the historical period with and without the policies - including building codes - in order to make the comparison (Marshall and Gorin, 2007).

Finally, it is well-known among experts that compliance with existing building codes is problematic; as a recent study noted

"Despite the lack of definitive national-level studies regarding building energy code compliance, and existing state studies which are difficult to compare and contrast, the available data signals a significant and widespread lack of compliance" (Building Codes Assistance Project, 2008).

These considerations indicate the value of an alternative, statistical approach to retrospective estimation of savings resulting from building codes. In this paper, by incorporating the adoption of these regulations at the state level in different starting years, we attempt to identify their impact on aggregate demand. Further, adding these regulations into the estimation will potentially improve the efficiency of the demand estimation.

Chapter 10

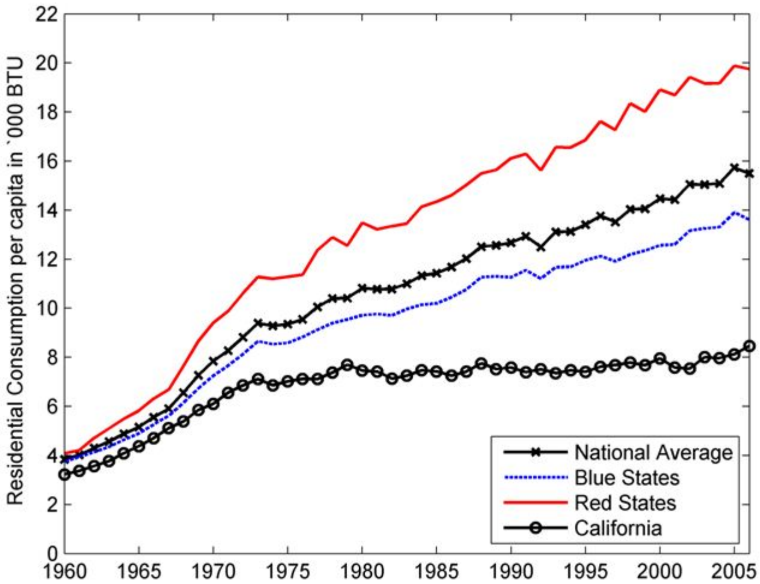
Data

10.1 Electricity Data

For each state and year we observe annual total electricity consumption for the residential sector in British Thermal Units (BTUs) from the Energy Information Administration's State Energy Data System (EIA, 2009). The database covers the years 1960 through 2006 for the 48 continental states. In order to translate total consumption into per capita consumption we obtain state level population estimates from the Bureau of Economic Analysis (BEA, 2009) for the same period. Figure 10.1 displays trends in the per capita series. The national average displays continued growth throughout the entire period. There is a noticeable slowdown in the growth rate in the early 1970s. If we split the states by political preferences, using the most recent presidential election as a guide, we can examine the differential trends in "blue" versus "red" states. Both series display a break in the trend in the early 1970s, yet the leveling off is much stronger in the blue states. If we look at California separately, the picture displays the often cited zero growth in per capita electricity consumption since 1974, which is often called the "Rosenfeld Curve". In addition to quantity consumed, we observe the average price of electricity for the residential sector as well as the average price of the main substitute source of energy in the sector, natural gas from 1970 on. The fact that we only observe the average price, instead of the marginal price, results in the price variable being endogenous in the empirical model.

Table 10.1 displays the summary statistics. The first four columns of numbers show within state and overall variation in each of the variables. Per capita electricity consumption displays a significant degree of within state as well as overall variation. The last four columns of table 10.1 display the summary statistics for states which adopted a building code at any point in the sample versus states that never did. The control states have a slightly higher average consumption and lower electricity and natural gas prices. This difference in prices alone makes it necessary to control for these confounders in order to obtain a consistent estimate of the effect of building codes on per capita electricity consumption.

Figure 10.1: Per Capita Residential Electricity Consumption Trends.



Note: Figure depicts the population weighted average per capita electricity consumption in BTU for the states with a majority voting for the democratic candidate for president in the 2008 presidential election (blue states), and the republican candidate (blue states) as well as the national average and California. Source: EIA State Energy Data System (2009)

Table 10.1: Summary Statistics

Sample		Complete (n=1,776)				Treated (n=1,591)		Control (n = 185)	
Variable	**Variation**	**Mean**	**S.D.**	**Min.**	**Max.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Electricity Consumption	*overall*	12.88	4.18	4.76	23.99	12.86	4.25	13.03	3.47
	within		2.55	5.24	20.38		2.20		2.83
Electricity Price	*overall*	0.15	0.10	0.01	0.58	0.15	0.10	0.14	0.08
	within		0.10	0.07	0.52		0.10		0.08
Natural Gas Price	*overall*	0.05	0.04	0.00	0.24	0.05	0.04	0.04	0.04
	within		0.04	0.02	0.22		0.04		0.04
Per Capita Income	*overall*	23.21	5.13	11.48	44.20	23.33	5.17	22.13	4.71
	within		4.04	13.48	36.02		4.09		3.57
Cooling Degree Days	*overall*	1.09	0.78	0.08	3.88	1.06	0.80	1.31	0.05
	within		0.14	0.59	1.64		0.14		0.16
Heating Degree Days	*overall*	5.24	2.04	0.40	10.75	5.24	2.08	5.25	1.73
	within		0.37	3.72	6.75		0.36		0.41
New Housing Stock (per capita)	*overall*	6.05	4.34	0.14	27.34	5.70	4.11	9.05	5.03
	within		2.70	12.18	14.58		2.41		4.48
New Housing Stock * Building Code	*overall*	4.72	5.92	0.00	31.37	5.27	6.02	N/A	N/A
(per capita)	*within*		4.46	10.36	21.01		4.72		N/A
Building Code Intensity	*overall*	0.31	0.34	0.00	1.00	0.35	0.34	N/A	N/A
	within		0.22	0.58	1.02		0.23		N/A

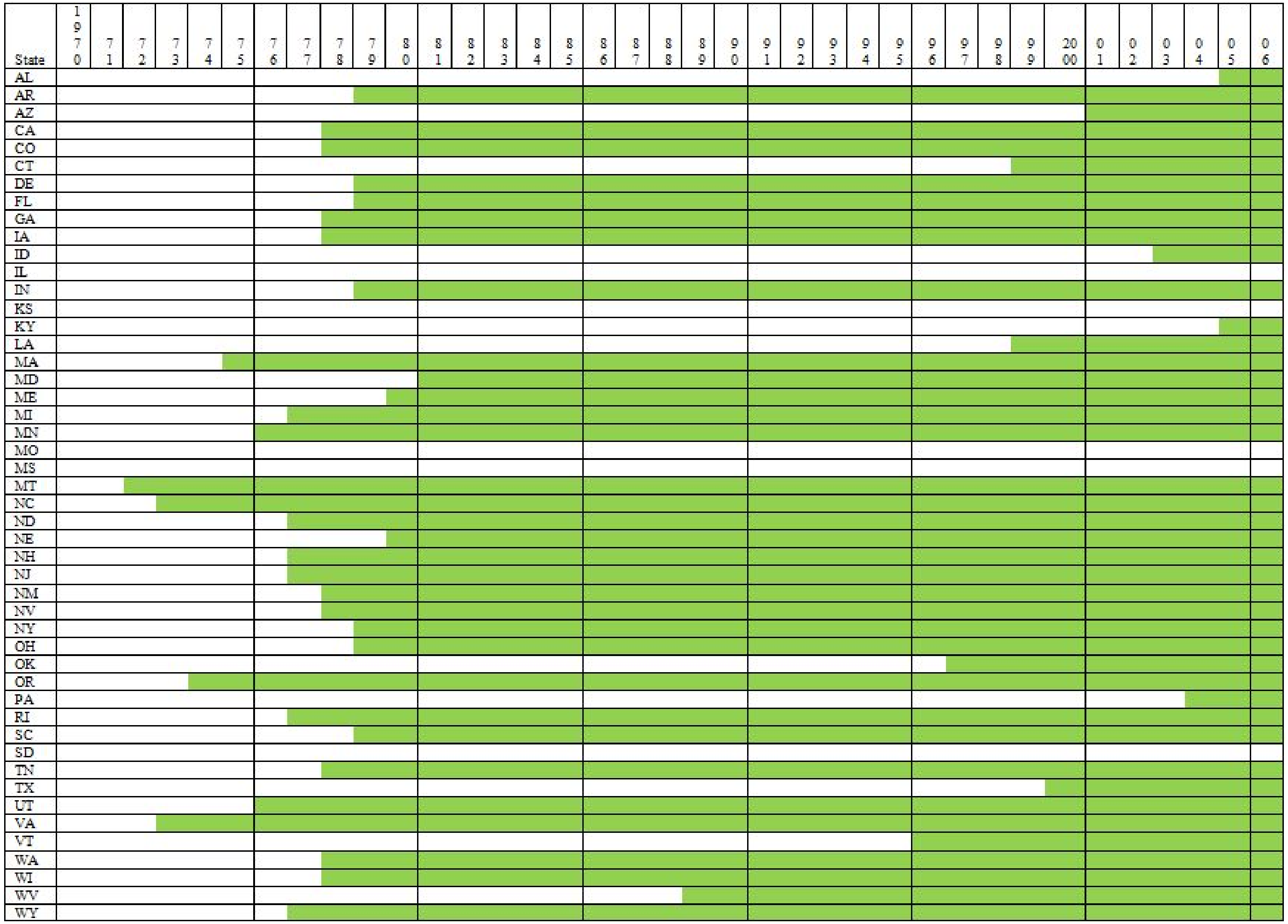
Note: T-test fails to reject the null of equality of means between treatment and control states in the pre-treatment year 1970 for all variables.

10.2 Building Code Data

We obtained data on the adoption and implementation of building codes at the state level from the building codes assistance program (BCAP, 2009). BCAP is a joint initiative of the Alliance to Save Energy (ASE), the American Council for an Energy-Efficient Economy (ACEEE), and the Natural Resources Defense Council (NRDC). It is partially funded by the US Department of Energy (DOE) and the Energy Foundation. BCAP assists state and local regulatory and legislative bodies with custom-tailored assistance on building energy code adoption and implementation. BCAP has a database, which contains detailed information on the current status of state level commercial and residential building codes as well as their history. We use the history section of the building codes database to construct a date of first implementation of building codes at the state level. In order to confirm the accuracy of these dates, we cross checked with state agency web sites to confirm that the dates are indeed implementation dates and not adoption dates. Figure 10.2 displays the building code dates for each state.

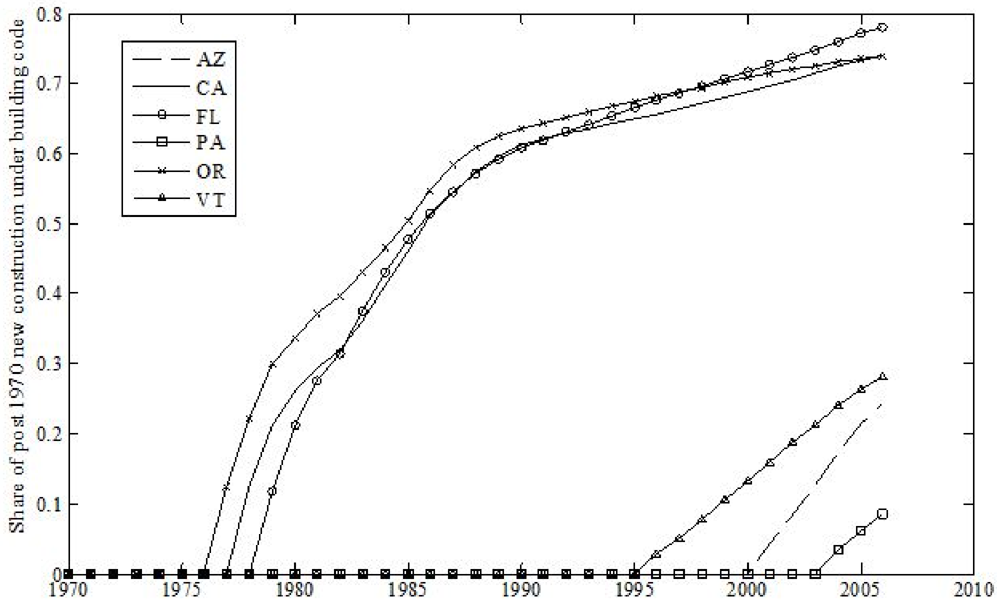
One could be tempted to use the binary indicator of whether state had a building code or not in a given year as the policy variable. This measure, however, would have several drawbacks. First, it ignores the heterogeneity in intensity of treatment across states. Since only new buildings and additions to existing buildings are subject to building code restrictions, states with higher growth rates new construction are likely to see bigger savings form these policies. Second, building codes vary across states in their stringency and enforcement. Using an undifferentiated binary measure allows one to estimate an average treatment effect of the average policy, but glosses over potentially important sub-national policy variation. Finally, while we know the year of implementation there may be some error as to when the codes actually started being enforced on the ground. This measurement error leads to attenuation bias towards zero of our estimated coefficients.

Figure 10.2: Implementation of Building Codes by State.



The "binary" strategy to estimate the effect of building codes on electricity consumption is to simply do a comparison in means before and after in treated versus control states controlling for other confounders. This difference in difference strategy is a valid identification strategy if one has random assignment in treatment. Since building codes apply to new construction and remodeled existing construction only, this strategy glosses over the fact that states with a higher rate of new construction are likely to see a bigger effect of building codes. We have therefore partially hand coded new building permits at the state level from a set of Census Bureau sources (US Census, 2009), to arrive at a measure of new construction before and after the implementation of a building code in a given state. While building permits are not a perfect measure of new construction, they likely are a good proxy and are comparable across states. The empirical measure we use to identify the effect of building codes on per capita residential electricity consumption is the share of the stock of new construction since 1970 which was conducted while a building code was in place. Figure 10.3 displays our policy measure for six selected states and displays significant variation across these states.

Figure 10.3: Share of New Construction Permitted Under Building Code.



Note: The figure depicts the share of housing permits issued after the passing of a state specific building code in the total stock of building codes issued since 1970.

ACEEE (2008) and Horowitz (2007) correctly point out that building codes vary in both their stringency and degree of enforcement across states. In order to explore the potential impact of heterogeneity in building code intensity (e.g. stringency and enforcement), we have collected ACEEE's indicator of building code stringency and enforcement. ACEEE (2008) reports a score for compliance for residential and commercial building codes ranging from one to five and an enforcement score, which ranges from zero to three. We exclude the commercial score in order to arrive at an overall score which we scale to range from zero to one for states with residential building codes. Our measure of intensity for a given state and year is defined as the product of this intensity indicator multiplied by a dummy whether the state had a building code during a given year or not.

10.3 Other Data

The remaining major confounders of residential electricity consumption found in the empirical literature are income and weather. We have obtained per capita personal income for each state back to 1970 and converted it into constant year 2000 US$ using the national CPI. This measure of income is a standard measure used in state panel data studies, since gross state product is not available as a consistent series back to 1970 due to the switch from the SIC to NICS product accounting system.

Finally we have obtained annual measures of weather relevant to heating and cooling demand in the form of heating and cooling degree days at the state level. Degree days are quantitative indices designed to reflect the demand for energy needed to heat a home or business and are non-linear in temperature (NOAA, 2009). They report population weighted degree days using decadal census information to re-weight degree days. A smoother measure of degree days would entail calculating the weights on an annual basis, but unfortunately such a measure is currently not available. Another potential problem is that the base temperature, which is currently set at 65° F for CDD and HDD should be different for different areas, yet this is likely a minor issue.

Chapter 11

Empirical Model

In order to estimate the effect of building codes on per residential electricity consumption, we estimate the following equation:

$$\log(q_{it}) = \beta_1 p^e_{it} + \beta_2 p^{ng}_{it} + \beta_3 y_{it} + \beta_4 CDD_{it} + \beta_5 HDD_{it} + \beta_6 Share_{it} + Z_{it}\gamma + \varepsilon_{it} \quad (11.1)$$

where is q_{it} is state i's per capita residential electricity consumption (million BTUs per person) in year t. p^e_{it} is the real average price of electricity to residential customers ($ per million BTUs), p^{ng}_{it} is the real average price of natural gas to residential customers ($ per million BTUs). y_{it} is real per capita income (thousands dollars), CDD_{it} and HDD_{it} are cooling and heating degree days respectively. $Share_{it}$ is the share of total new construction since 1970 which was permitted while a building code was active in state i. Z_{it} is a vector of random or deterministic variables which vary across states or time or both. We will estimate equation (11.1) using panel data from 48 U.S. states excluding Hawaii and Alaska. The sample period is from 1970 - 2006, since prices are only available from 1970 on.

One issue that has been widely pointed out in electricity demand estimation is price endogeneity due to the increasing block price structure in place in most states (e.g. Hanemann 1984). The average price will be affected by the quantity consumed, which leads to an upward-bias (in absolute value) in the estimate of the demand response. To deal with this simultaneity problem, a common procedure is to use estimated marginal price data instead (Berndt, 1996), which unfortunately are not available for the 48 states and 37 years of data in our sample. It is often noted that households do not know the marginal price of electricity and respond to the total monthly energy bill. Marvin (2004) provides supporting evidence of such a price exogeneity assumption. While Baltagi et al. (2002) and Maddala et al. (1997) in similar studies assume price exogeneity citing these studies, we instrument for price using predetermined lagged values of average price.

There are three different avenues to estimate the parameters for equation (11.1). First, the most basic approach is to assume homogeneity of the parameters across states and estimate a single demand equation by pooling data. Ordinary Least Squares is consistent if the disturbance is orthogonal to all right hand side variables. Second, one can allow for a limited degree of heterogeneity in time invariant unobservables by adopting a fixed effects estimator. This approach still assumes that all coefficients, with the exception of the intercept, are identical across states. Finally, one can assume that all coefficients vary across states. Under this assumption, one could estimate the equations state by state, which results in very imprecisely estimated coefficients due to the short time series for any given state. One approach proposed in the literature to mitigate this effect is a so called "shrinkage estimator", that allows for some, but not complete, heterogeneity of the parameters by shrinking the coefficient from each state towards the overall mean. Maddala et al.(1997) tested and rejected the null hypothesis of homogenous coefficients in electricity demand estimation using an older version of the dataset employed in this paper. They show that when estimated state by state, the estimated coefficients have unreasonable signs. They argue that the shrinkage estimators are the most reasonable method.

However, Baltagi and Griffin (1997) explored a much larger number of estimators, including an instrumental variables estimator, and compared the plausibility and forecasting performance of these estimators using dynamic demand for gasoline in 18 OECD countries. They found that the homogeneous pooled estimators yielded much more plausible estimates and gave a better out-of-sample forecast. More recently, using the same data employed in this paper, albeit for a shorter time period, Baltagi et al. (2002) showed that the pooled estimators significantly outperformed the heterogeneous coefficient models for both U.S. demand for electricity and natural gas. Following Baltagi et al. (2002), we will assume homogeneity of the slope coefficients. The goal of the paper is to identify the effect of building codes adopted by states on residential per capita electricity consumption. We will control for time invariant heterogeneity across states by including fixed effects in our preferred estimation.

The most restrictive set of assumptions are that $\varepsilon_{it} \sim iid(0, \sigma^2)$, all independent variables are independent of this disturbance term and that there are no time invariant differences in unobservable across states or common unobservable shocks across states in any given year. As mentioned above, if these assumptions hold, ordinary least squares is consistent. Using OLS, the regulation treatment effect would be identified if the regulation variables were exogenous and thus not correlated with the error terms (ε_{it}). However, each state can choose whether or not to adopt regulations and that choice could be correlated with time invariant differences in unobservable characteristics of each state, which we do not include in the model. We partially control for this problem by assuming two-way error component disturbances: $\varepsilon_{it} = \alpha_i + \phi_t + u_{it}$ where $u_{it} \sim iid(0, \sigma^2)$

This specification allows us to capture the unobservable time invariant state specific effects (α_i) (e.g. unobservable geographic characteristics) that might be correlated with the regulation variables. One example where omitting fixed effects would result in bias of the regulation variables is that the median voter in adopter states may be more concerned about energy and environmental issues than the median voter in non-adopter states. These green voters are likely to consume less electricity in the absence of regulation. Without controlling for state fixed effects, this green voter effect wound confound the estimated effect of the policy. By including unobservable time effects (ϕ_t), we can control for the time-specific shocks that commonly affect per capita consumption in all states, such as oil shocks, recessions and federal policies applicable in all states.

If regulations were randomly assigned to states at the same point in time, we could identify the average treatment effect of regulations via a simple dummy variable approach. In that case, the average treatment effect would be measured by the difference in electricity consumption of the treatment group (states that adopt) and control group conditional on other state characteristics such climate, prices etc. Here, some states choose to adopt regulations while others do not, and states also adopt regulations at different points in time. Further, intensity of treatment varies by the degree of new construction in a given state as well as the stringency of the actual building code and its enforcement. A building code in a state with no new construction, a weak building code or any building code without enforcement will be ineffective. To address the first concern, we use the interaction of a dummy of building code treatment with the share of the total building stock constructed since 1970 (the first year of our sample) as our variable of treatment. This variable in theory can range from zero to one.

Since this variable is potentially endogenous, we use an instrumental variables technique to obtain consistent estimates of the effects of building codes on per capita residential electricity consumption. Since harsh winters in the early 1970s were the trigger for the first building codes, we instrument with twice lagged cooling degree days as well as lagged heating degree days. We also use the predetermined lagged share variables as instruments.

In order to explore the potential impact of heterogeneity in building code intensity across states, we estimate the following model:

$$\begin{aligned} \log(q_{it}) \quad = \quad & \beta_1 p^e_{it} + \beta_2 p^{ng}_{it} + \beta_3 y_{it} + \beta_4 CDD_{it} + \beta_5 HDD_{it} + \beta_6 Share_{it} \\ & + \beta_7 Intensity_{it} + Z_{it}\gamma + \varepsilon_{it} \qquad (11.2) \end{aligned}$$

where $Intensity_{it}$ is the building code intensity measure based on ACEEE (2008), which ranges from zero to one for each state-year with an active building code and is zero for all state/years without building codes. β_7 therefore measures the effect of a more intense building code. We would expect states with more stringent and well enforced building codes to have lower per capita residential electricity consumption.

Using a panel data set which includes states from both treated an control groups as well as both time periods (before and after treatment) and econometric techniques which control for factors leading to time differences in adoption and intensity of treatment, we hope to address the counterfactual question of how per capita electricity consumption would have changed if an adopting state did not adopt regulations.

Chapter 12

Estimation Results

Models (1) through (8) in table 12.1 show the results from estimating variants of equation (11.1). The first model is the pooled specification without building codes or fixed effects. The estimated coefficients should be interpreted as reflecting the short run response of electricity consumption in prices, climate and income. Since the estimated specification is a log-linear functional form the coefficients should be interpreted as the approximate percent change in residential per capita electricity consumption due to a one unit change in the covariate. In order to get an elasticity one multiplies the coefficient by a value of interest of the covariate (e.g. its sample mean). For the pooled model in column (1), the own price elasticity at the sample mean is -0.22, the cross price elasticity with natural gas is 0.35, the income elasticity is -0.11 and the elasticities for cooling an heating degree days are 0.17 and 0.05 respectively. All coefficients lie within the range of those found in the literature (e.g. Maddala et al. (1997)), with the exception of the negative income elasticity.

Column (2) controls for year and state fixed effects. The own price elasticity is now a smaller (-0.14), the cross price elasticity is -0.23 and the income elasticity is now a positive and significant 0.35. The CDD and HDD elasticities changed significantly, since we switch to an identification strategy relying on within state variation. The coefficient on CDD is no longer significant and the elasticity is 0.05, while the HDD elasticity is now a significant 0.16.

Column (3) addresses the issue pointed out in the previous section, that due to the fact that we observe average and not marginal prices, the price of electricity is endogenous. We therefore instrument with the predetermined lagged price of electricity and price of natural gas. The own price elasticity recovers slightly to -0.18, as we would expect a least squares coefficient being biased towards zero. The cross price elasticity becomes even smaller (0.09), which is consistent with the findings in the literature. The income elasticity is closer to other short run estimates found in the literature (0.10) as are the statistically significant coefficients on HDD and CDD.

Column (4) is the first model which includes our measure of policy - the share of newly permitted construction since 1970 under an active building code. The variable itself varies between 0, for states which either did not have a building code in a given year or did not have any construction after the implementation of the regulation, and one. The coefficient estimates on the remaining confounders are almost identical to those in model (3). The point estimate on the share variable is -0.053, which is suggests that if all construction in a given state has been built under an active building code, per capita electricity consumption is approximately 5% lower than in a state without such a building code.

Before we discuss what the magnitude of this coefficient implies, we want to check its robustness. One reason for potentially obtaining a significant and negative coefficient estimate on our policy variable is that states which have adopted building codes and have experienced significant new construction, may have had preexisting trends in per capita electricity consumption, which have nothing to do with policy, but may give rise to this significant coefficient estimate. We include linear time trends separately for states that have and have not adopted building code regulation to control for this potential phenomenon in column (5). The coefficients are almost identical to those in column (4). Model (6) includes second degree polynomial trends for building code and non building code states separately and again, the coefficients are almost identical.

Column (7) attempts to deal with the issue of endogeneity of our policy variable. Both the regulation and building construction are highly correlated with climate. Costs of construction are higher in years with sever climate outcomes (longer winters delay the construction season). Further, the first building codes were motivated by the severe winters in the early 1970s. We therefore use the first and second lag of HDD and CDD as instruments as well as the predetermined share variables. The coefficient on the policy variable moves very slightly away from zero, which is what one would expect. The other coefficient estimates, again, are almost identical to those in the previous columns.

Table 12.1: Ordinary Least Squares and Instrumental Variables Fixed Effects Regression Results

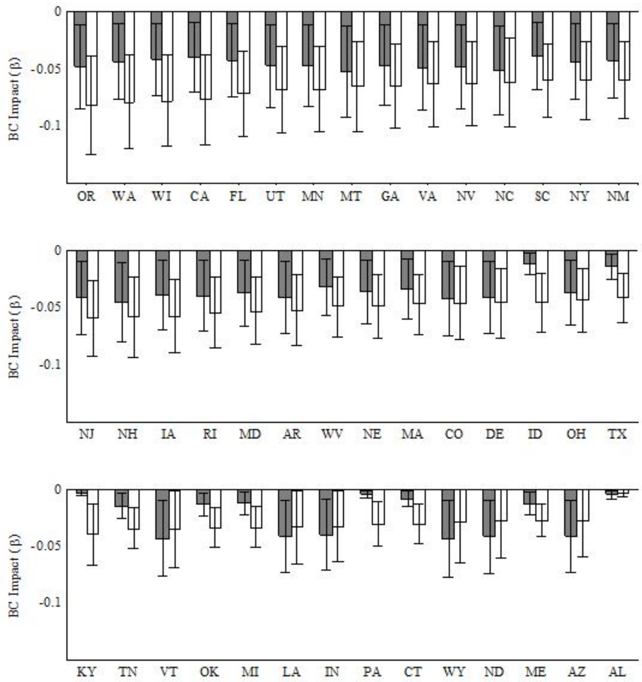
	Dependent Variable: logarithm of per capita residential electricity consumption								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Electricity Price	-1.480***	-0.942***	-1.220***	-1.207***	-1.209***	-1.208***	-1.188***	-1.185***	-1.174***
	(0.155)	(0.325)	(0.105)	(0.105)	(0.105)	(0.105)	(0.104)	(0.105)	(0.104)
Natural Gas Price	7.901***	4.819***	2.001***	1.985***	1.987***	1.985***	1.999***	1.996***	2.003***
	(0.335)	(0.782)	(0.250)	(0.250)	(0.250)	(0.250)	(0.246)	(0.249)	(0.245)
Per Capita Income	-0.004*	0.015**	0.004*	0.005*	0.005*	0.005*	0.005*	0.005*	0.005*
	(0.002)	(0.006)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Cooling Degree Days	0.154***	0.050	0.084***	0.084***	0.084***	0.084***	0.088***	0.088***	0.090***
	(0.018)	(0.040)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)
Heating Degree Days	0.010	0.031**	0.051***	0.049***	0.049***	0.049***	0.046***	0.045***	0.045***
	(0.007)	(0.013)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
Building Code Construction Share				-0.053**	-0.052**	-0.052**	-0.056**	-0.056**	-0.037**
				(0.020)	(0.020)	(0.020)	(0.021)	(0.021)	(0.022)
Building Code Intensity									-0.050***
									(0.019)
State Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV Electricity Price	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BC/Non-BC Trends Linear	No	No	No	No	Yes	Yes	Yes	Yes	Yes
BC/Non-BC Trends Non-Linear	No	No	No	No	No	Yes	Yes	Yes	Yes
IV BC Share	No	No	No	No	No	No	Yes	Yes	Yes
Appliance Standard Trend & Dummy	No	No	No	No	No	No	No	Yes	Yes
No. of Obs.	1,776	1,776	1,728	1,728	1,728	1,728	1,680	1,680	1,680
R^2 **(within)**	0.378	0.799	0.826	0.828	0.828	0.828	0.819	0.819	0.819
Number of States	48	48	48	48	48	48	48	48	48

Note: Robust standard errors are in parentheses (***$p<0.01$, **$p<0.05$, *$p<0.1$)

Finally, five states had adopted appliance standards prior to the federal appliance standards passed in 1987. If these states are also building code adopters and appliance standards are actually effective at reducing electricity use, we may confound the impact of building codes with that of state specific appliance standards. For building codes we do have a proxy for intensity of treatment, which is the amount of new construction. We do not know how many air conditioners and refrigerators satisfying the state specific appliance standards were sold in any given year. We therefore control for appliance standards via a dummy for whether the standard was on or off in any given year. We also include a trend variable, which is 0 if a given state does not have an appliance standard in a given year, 1 for the first year of the appliance standard, 2 for the second year etc. Column (8) displays the estimation results from this model. Again, the coefficient estimates on the building code policy variable remains roughly unchanged. The other coefficients are also almost identical to the previous specifications. While we do not display the estimates on the appliance standard variables here, the coefficient on the dummy variable is not statistically different from zero, and the coefficient on the trend variable is very close to zero, albeit statistically different from zero.

Column (9) augments the model from column (8) by including the building code intensity measure. As expected, the coefficient estimates on the building code policy variable of interest ($share_{it}$) moves slightly towards zero and is only statistically different from zero at the 10% level. The other coefficients are also almost identical to the previous specifications. The measure of building code intensity is statistically and economically significant. A state with the most stringent implementation of their building codes, such as California or Oregon, can expect to have a 5% lower per capita electricity consumption relative to states with a zero rating. The overall effect of a building code for a given state and year can therefore be calculated by $\beta_6 \cdot Share_{it} + \beta_7 \cdot Intensity_{it}$.

Figure 12.1: Specific impact of building codes for the year 2006.



Notes: The grey bars indicate the predicted impacts of building codes from model (8) from table 12.1. They are obtained by multiplying the 2006 building code construction share variable times the estimated coefficient. The whiskers indicate the 95% confidence interval. The white bars indicate the predicted impacts of building codes from model (9) from table 12.1. They are obtained by multiplying the 2006 building code construction share variable times the estimated coefficient and adding the product of the building code intensity value for 2006 with its estimated coefficient. The whiskers indicate the 95% confidence interval.

We now turn to putting the coefficient estimates on the share variable into perspective. Figure 12.1 displays the estimated impact of state building codes from both models (8) and (9). The grey bars show the effect building codes on per capita electricity consumption based on model (8). For each state we calculate $\beta_6 \cdot Share_{it}$ and its 95% confidence interval, as indicated by the whiskers.

The white bars show the effect building codes on per capita electricity consumption based on model (9). For each state we calculate $\beta_6 \cdot Share_{it} + \beta_7 \cdot Intensity_{it}$ and its 95% confidence interval, as indicated by the whiskers. Figure 12.2 displays a map of the share variable for the year 2006. States with the majority of the new construction activity after the implementation of the building code appear as dark green. It is not surprising that early adopter states experiencing recent rapid population growth such as Nevada, Georgia and Utah appear as states with a high share variable here. This variable drives the estimated effects in figure 12.1 of the grey bars. Once we combine these effects with the building code stringency variable, we obtain large estimated effects for Oregon, Washington, Wisconsin, Florida and California, due to their large share of new construction as indicated by figure 12.2, their long history of stringent and well enforced building codes. It is important to note, however, that if one is interested in aggregate savings, one has to consider the total number of new buildings in a given state. Figure 12.3 displays the total number of new building permits by state, which shows that California and Florida due to their size and growing populations account for 28% of the building permits issued in states with an active building code.

Figure 12.2: Year 2006 Share of post-1970 permitted new construction conducted under an active building code by state.

It is interesting to compare our estimates of savings from building codes to estimates obtained from bottom up engineering models. While we are not aware of a comprehensive accounting exercise using the most recent set of building codes at the national level, we have obtained such estimates from the California Energy Commission (CEC, 2007). By calculating attributed savings from each type of policy across California's utilities, the CEC arrives at 5,259 GWh in savings due to building codes out of a total consumption of 86,069 Gwh resulting in attributed savings of roughly 6%. Our estimate for the same year for model (9) is 7.7%. Due to the sizable confidence interval, it includes the estimate provided by the CEC. The estimate based on model (8) is savings of 4.03%, yet the confidence interval still includes the figure cited by the CEC. The smaller estimate of model (8) is the average effect of building codes

Figure 12.3: Total number of building permits issued for each state since enactment of building codes.

across states. If *e.g.* California, has a more stringent or better enforced building code, we would expect that savings for that specific state are higher. The identification strategy in this model, however, cannot reliably identify state level treatment effects, since it uses across state and time variation in policy introduction and building intensity to identify the average policy effect. Model (9) attempts to overcome this shortcoming and identifies a much larger effect for states with more intensive building codes. Further, our estimate does not control for spillover benefits from first adopter states. For example, the California standards for refrigerators are argued to have had a US wide effect long before federal standards were promulgated in 1987. Our estimates therefore only capture the effects of the policy on the treated.

One interesting question is how much energy in 2006 has been saved overall from the residential building codes currently in place? In order to calculate this figure, we set the building code indicator to zero for all states and calculate overall energy consumption. If we calculate total predicted savings for each state and predict total savings from model (8), we arrive at an estimate of 2.09% in reductions of aggregate residential electricity demand due to these programs for the year 2006 counting all states in the denominator. The model allowing for heterogeneity in building codes estimates overall savings from building codes using model (9) at 4.98%. Our interpretation is that the true number lies in between these two figures, since the estimate on the building code stringency variable is likely to be confounded with the effect of other programs. A regulator imposing a more stringent building code is also likely to impose other more aggressive conservation measures, which may be captured by our intensity measure.

Chapter 13

Conclusion

We regard this paper as a first step in using an econometric model to identify an average treatment effect of the average building code on residential electricity consumption. We use information from time varying state specific regulation adoption to identify the effect of the regulations on consumption. We find a significant effect of building codes on residential per capita electricity consumption ranging from 0.3-5% depending on the state. Aggregating savings to the national level, our estimates suggest savings in residential electricity consumption of 2.09-4.98% for the year 2006. Various studies have shown that compliance with building energy codes may be low. According to a ACEEE (2003) report, in most cases, builders are not completely compliant with energy codes. This may be due to a lack of informed builders, the complexity of the building code and insufficient training of code officials. Our results show that if authorities were able to ramp up compliance and enforcement, current estimates of program effectiveness likely represent a lower bound of what is possible. Further research is necessary to quantify the range of state specific treatment effects of building codes on residential electricity consumption. Another interesting topic of research is to assess the costs of these building codes and compare them to the derived benefits.

Part III

The Impact of Ozone on Crop Yields in the U.S.: Evidence from Corn and Soybeans

13.1 Introduction

In contrast to ozone in the stratosphere which protects life on Earth from ultraviolet (UV) radiation, ground-level ozone (ozone hereafter) is harmful to breathe and causes respiratory problems such as congestion, asthma and lung damage (EPA, 2006). USDA (2008) states that ozone also causes more damage to crops, trees and other vegetation than all other air pollutants combined. EPA (2006) estimates that ozone causes $1 billion annually in crop losses in the United States. Realizing the significance of ozone's adverse effects, the U.S. includes ozone as one of the six criteria air pollutants regulated by National Ambient Air Quality Standards (NAAQS) under the Clean Air Act. Two types of national air quality standards for ozone were set: the primary standard that aims to protect human health and the secondary standard that aims to protect public welfare including protection against damages to animals and crops. Since 1997, these two standards had been set at the 0.08 part per million (ppm) for 8-hour ozone standard and it was tightened to 0.075 ppm in 2008.

Ozone is formed by an interaction of Nitrogen oxide (NOx), volatile organic compounds (VOCs) and sunlight. The primary sources of NOx are motor vehicles and electric power plants[1]. VOC emissions come from the use of products such as fuels, paints and solvents. The main source of VOCs, however, are bioganic. Trees, shrubs and plants release VOCs, which combine with NOx and sunlight to form ozone. Rural areas therefore are NOx limited, resultding in reductions in ozone due to reductions in NOx. Ozone can be transported hundreds miles away by wind from the source of its precursors especially in the rural areas where crops are grown.

Ozone can cause negative effects on a number of plant processes, including photosynthesis, water use efficiency, dry matter production, flowering, pollen tube extension, and thus yield (Krupa, 1997, Mackee et al., 1997). To investigate the negative effects of ozone on crop yields, most of this research utilized open-top field chambers experiments (Mauzerall and Wang, 2001). During 1980-1986, the National Crop Loss Assessment Network (NCLAN) conducted a series of open-top chamber experiments and found that yield loss per seasonal 7-hour per day mean ozone concentration ranges from 1 to 30 percent varying by crop type and location (Heagle, 1989).

The major drawback of open-top chamber studies is that these controlled chamber experiments may not accurately reflect actual pollution impacts under field condition. As Leung et. al (1982) note, "...damage functions that are derived from laboratory or controlled experiments are not necessarily correlated closely with actual farm situations." A chamber usually increases temperature, decreases light intensity and decreases precipitation. In addition, the filtered chambers might also remove some nutrients from air that plants need for growth and remove some toxic substances other than ozone. Further, these experiments do not account for offsetting actions taken by famers. These factors could affect crop growth and thus bias the estimated

[1]http://www.epa.gov/air/ozonepollution/basic.html

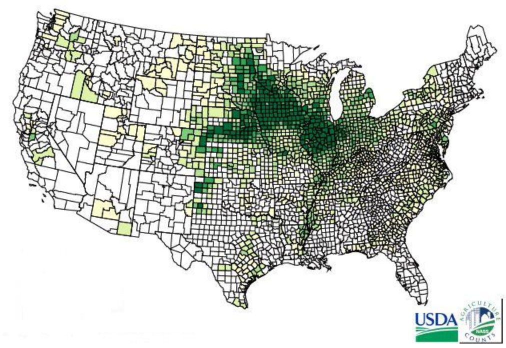


Figure 13.1: Corn production in 2008

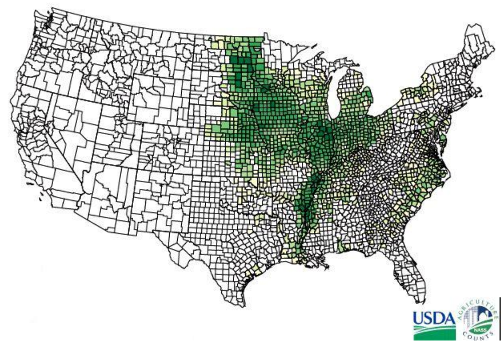


Figure 13.2: Soybeans production in 2008

Note: These two figures show the productions of corn and soybeans by county in 2008 respectively. The shaded areas show where corn and soybeans are grown and the darker shade represents the higher level of production.

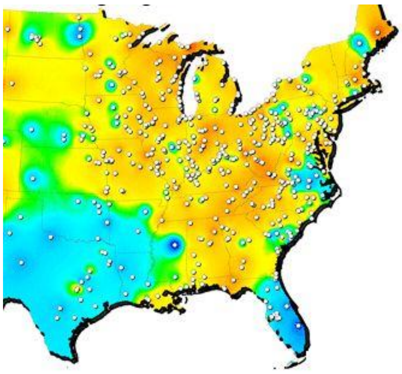
impact of ozone on crop yields.

An alternative approach is to use actual field production data to econometrically estimate the impacts of ozone on crop yields. Only a few studies using this approach have been done (Leung et al.(1982), Garcia et. al (1986), Rowe and Chesnut (1985) and Westenbarger and Frisvold (1995)). Using regional data and simple regression, they also found a significant adverse impact of ozone on crop yields similar to the NCLAN results. None of these papers; however, used nation-wide panel data and a fixed effect model to reduce the bias from unobservable covariates such as soil quality.

Corn and soybeans are the two major crops grown in the U.S. In 2008, the total planted area of the two crops combined was approximately 160 million acres and the value of their production is about 80 billion dollars (USDA, 2009). Figures 13.1 and 13.2 show the production of corn and soybeans by county in 2008 respectively. The shaded areas show where corn and soybeans are grown and the darker shade represents the higher level of output. These two crops are heavily produced in the Midwest and the Northeast where coal power plants are also located as shown in figure 13.3. During summer months when demand for electricity is high, coal-fired power plants emit a high level of NOx. When NOx combines with sunlight, ozone is formed and the crops grown in that region and downwind are prone to be damaged by ozone. Figure 13.4 shows monthly average ozone concentration across monitors located in agricultural land during 2000 to 2006. The two vertical lines indicate the beginning and the end of a growing season (March to August). The mean ozone reaches its peak around May to August.

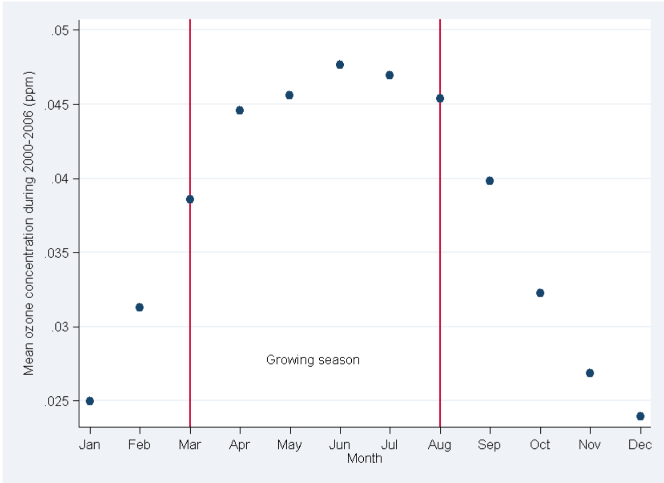
In an effort to reduce ozone in the eastern U.S., the NOx SIP call program requires 22 eastern states to put NOx reduction measure in place by 2003. Under the NOx SIP Call program, the NOx Budget Trading Program (NBP), a market-based cap and

Figure 13.3: Coal power plants in 2006



Note: The dots in this figure indicates locations of coal power plants in 2006. Most of the coal power plants are located in crop cultivated regions.

Figure 13.4: Monthly mean ozone concentration during 2000-2006



trade program is established to reduce NOx emissions during the summer months, when ozone concentrations are high. Compliance with the NOx SIP begun on May 1, 2003. According to EPA (2008), ozone season NOx emissions have decreased by more than 60% between 2003 and 2008. Butler and Lee (2007) estimated that the NBP reduces 8-hour ozone concentration on average by 10% below the level seen before the program.

In this paper, I construct a U.S. county-level panel data of crop yields and ozone concentrations from 1990 to 2006. To my knowledge this is the first paper that uses nation-wide panel data to econometrically estimate the impact of ozone on crop yields. This paper focuses on the impact of corn and soybeans because they are two major crops grown heavily in the ozone exposure regions. In an attempt to reduce omitted variable bias, the fixed effect model is used to reduce the bias from time-invariant unobersavable covariates. I also instrument ozone concentrations with a state's NBP compliance year dummy. I find that the elasticity of crop yield with respect to various ozone measures range from -0.60 to -0.73 and -0.57 to -0.64 for soybeans and corn respectively. Back of the envelope estimate of crop losses in the year 2008 is about 5 billion dollars.

The remainder of the paper is organized as follows. The next chapter discusses the impact of ozone on crop yields. Chapter 15 describes the data. Chapter 16 presents the econometric model. Chapter 17 presents estimation results and chapter 18 concludes.

Chapter 14

Impact of ozone on crop yield

To investigate the negative effects of ozone on crop yields, most of this research utilized open-top field chambers experiments (Mauzerall and Wang, 2001). This method has been used in Europe, U.S. and Asia to assess ozone effects on a range of crops such as soybeans (Ficus et al.,1997), corn (Rudorff et al., 1996), wheat (Tiwari et al., 2005), cotton (Zlatev et al., 2001) and rice (Wahid et al., 1995, Maggs and Ashmore, 1998). They have found evidence of yield reduction ranging from negligible to over 30 percent, depending on the type of crop and ozone measure used. This method was also used by the most comprehensive study assessing the economic impacts of crop losses from air pollutants in the U.S. conducted by the National Crop Loss Assessment Network (NCLAN) from 1980-1986. In this study, field-grown crops were exposed to either non-filtered or filtered-air in open-top chambers. For 7 hour per day during growing season, various amounts of ozone were added to air pulled into the bottom of the non-filtered chamber or ozone is filtered out before entering the filtered chamber to keep concentrations at the targeted level. The yield-ozone dose responses are then recorded. NCLAN reports estimated percentage yield loss per various concentrations of seasonal 7-hour per day mean ozone ranging from 1 to 30 percent varying by type of crop and ozone locations[1] (Heagle, 1989).

An alternative approach is to use actual field production data to econometrically estimate the impacts of ozone on crop yields. Only a few studies using this approach have been done. Leung et al.(1982) used principal component analysis with county-level panel data from Southern California. The negative impact of mean ozone on the yield of 8 of the 9 crops analyzed in this study, varied from 0% (celery) to 57% (avocado) depending on crop and location[2]. Rowe and Chesnut (1985) used linear regression and found that estimated yield losses are close to those extimated with NCLAN chamber studies and the nonlinear specifications showed little or no im-

[1]For comprehensive list of dose response function see Table 3 of Heagle (1989).

[2]The crops are alfalfa, avocado, celery, lemon, lettuce, navel orange, valencia orange, strawberrry, and tomato. No yield reduction from ozone was found for celery.

provement and provided almost identical yield loss estimates[3]. Garcia et. al (1986) used farm-survey panel data in Illinois and found the elasticity of mean ozone (ppm) on grain farm profit to be -0.476. The most recent econometric study is done by Westenbarger and Frisvold (1995) using cross-section farm-level data for the eastern United States. They found that the mean elasticity of yield to ozone exposure is -0.19 and -0.54 for corn and soybeans respectively.

[3]The crops in this study are dry beans, cotton, grapes and potatoes

Chapter 15

Data

15.1 Ozone data

Daily ozone data at the monitor level from 1995 to 2006 are collected from the EPA's Air Quality Standards (AQS) database[1]. This database reports hourly air pollution readings from the EPA's network of air quality monitors nationwide. The data set also includes monitor latitude and longitude and type of land use at each monitor's location.

I limit the sample of data as follows. First, since I would like to estimate the effects of ozone on crop yield, I only keep the data from the monitors located in agricultural areas. Second, following the EPA data standard, I drop all monitor-days for which observations are not recorded for at least 9 hours between 9am to 9pm. I also drop monitor-years for which more than 25% of the days during the growing season (1 March to 31 August) report no observation. Finally, I only keep the data during growing season because crops are only exposed to ozone during the growing season.

There is no consensus on which ozone measure is the best representation of the potential response of crops to changing ozone concentrations over a growing season. In this paper, I construct three daily ozone concentration measures which will be used to calculate the seasonal ozone measure later on. They are (1) the daily maximum ozone, (2) the daily 8-hour maximum ozone and (3) the weighted daily mean ozone. The daily maximum ozone is the maximum of ozone during 9am-9pm. The daily 8-hour maximum ozone is calculated by taking the maximum of the 8-hour moving average of hourly ozone in each day during 9am to 9pm. These two measures will capture the impacts of high ozone concentration on crop yield.

Since low ozone concentrations are expected to have little impact on crop yield compared to the high ozone concentration, I calculate the weighted mean of ozone concentration rather than a sample mean. Lefohn and Runeckles (1987) suggest the

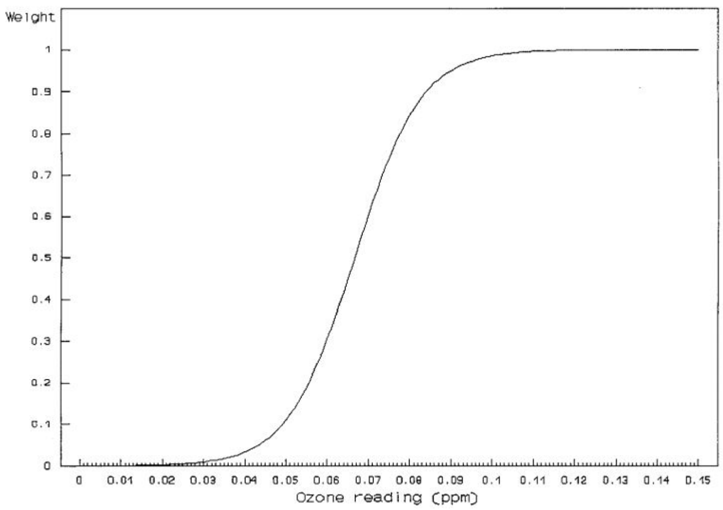
[1] http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm

cumulative exposure index (CEI) that weights each hourly ozone by the sigmoidal weighting scheme.

$$w_i = 11 + Me^{-Ac_i} \tag{15.1}$$

where w_i is the weighting factor, c_i is the ozone concentration, $A = 126$ and $M = 4403$. Figure 15.1 depicts the weight function. The weight is increasing at an increasing rate and then increasing at a decreasing rate when ozone concentration reaches 0.065 ppm. I weight each hourly ozone c_i with its corresponding w_i then the mean ozone is calculated to get weighted daily mean ozone.

Figure 15.1: Sigmoid weighting Function



15.2 Crop yield data

Corn and soybean yield data are collected from the United States Department of Agriculture's website[2]. This database reports annual crop yield data at the county level. A crop yield is defined production (bushels) divided by acres harvested in each county.

[2]http://www.nass.usda.gov/QuickStats/Create_County_All.jsp

15.3 Weather data

Weather data come from the National Climatic Data Center's Cooperative Station Data (NOAA), which provide daily minimum and maximum temperatures and rainfall at more than 20,000 weather stations across the United States. These weather stations are not typically located near a pollution monitor and many have missing observations. To obtain daily weather observation at each pollution monitor, I use the following algorithm by Auffhammer and Kellogg (2009). First, they calculate the Vincenty distance of each pollution monitor to all weather stations. They then identify the ten closest weather stations to each pollution monitor, provided that each is less than 50 miles from the monitor and the elevation difference between the monitor and the station is less than 500 vertical feet. Of these stations, they identify the "primary station" as the closest station for which 50% of the pollution monitor's daily readings can be matched to the station's weather data. They then match the climate variables for this station to the time series of ozone measurements.

Following the steps, 10.2% of the daily ozone measurements are not matched to a full set of weather variables from a primary station. These missing values are filled by first regressing, for observations in which the primary weather station was active, the relevant weather variable for the primary station onto the same variable for the remaining nine closest stations. The predicted values from that regression are used to replace the missing values. Following this step, primary station observations are still missing whenever one of the remaining nine closest stations is also missing an observation. To estimate the remaining missing values, the above step is repeated with the 8 closest stations, then the 7 closest stations, etc. At the end of this procedure, less than 0.1% of the remaining ozone monitor observations are still missing a matching climate observation. Those observations are then dropped.

For estimation purposes, the three ozone measures and weather data need to be converted to a dataset at the county-year level. To do so, the pollution monitor locations are mapped onto the countiy boundaries. If there are multiple monitors in a county, each daily ozone measure and each weather variable are averaged across monitors within the county. Then, I average each of the daily ozone measure and each weather variable over a growing season (March to August) to get the county-year panel data. Table 15.1 shows summary statistics from the pooled dataset over 1990 - 2006 by type of crop. The mean of soybean and corn yield are about 35 and 115 bushels per acre. The mean of the three ozone measures are similar ranging from 0.05 - 0.06 ppm. The maximum ozone could go up to 0.09 ppm and the maximum 8-hour ozone could be as high as 0.076 which is higher than the EPA standard.

Table 15.1: Summary statistics summary from the pooled dataset over 1990 -2006 by type of crop

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
			Soybeans		
Soybeans yield (bushels/acre)	34.833	9.248	8.9	58.8	1,495
Mean ozone (weighted)(ppm)	0.046	0.005	0.024	0.064	1,495
Maximum ozone (ppm)	0.057	0.007	0.031	0.082	1,495
Maximum 8-hour ozone (ppm)	0.052	0.006	0.028	0.071	1,495
Mean temperature (°F)	67.112	4.385	52.816	82.867	1,495
Maximum temperature (°F)	78.089	4.753	63.366	92.956	1,495
Minimum temperature (°F)	56.739	4.293	42.272	72.778	1,495
Total rainfall (mm)	2,097.909	606.239	0	5,099	1,495
			Corn		
Corn yield (bushels/acre)	115.714	32.531	17.1	208.9	1,919
Mean ozone (weighted) (ppm)	0.046	0.006	0.018	0.068	1,919
Maximum ozone (ppm)	0.058	0.007	0.027	0.090	1,919
Maximum 8-hour ozone (ppm)	0.052	0.006	0.021	0.076	1,919
Mean temperature (°F)	66.527	5.153	44.019	90.981	1,919
Maximum temperature (°F)	77.763	5.66	59.246	104.222	1,919
Minimum temperature (°F)	55.368	5.197	28.793	77.741	1,919
Total rainfall (mm)	1,887.182	755.823	0	5,099	1,919

Chapter 16

Empirical Strategy

Using panel data from 1995 to 2006, I identify the effect of ozone on crop yield using the following fixed effect model specification.

$$\ln(yield_{it}) = \alpha_0 + \beta_1 \ln(ozone_{it}) + \sum_{k=2}^{K} \beta_k x_{kit} + \theta_i + \epsilon_{it} \tag{16.1}$$

From equation (16.1), i indexes county and t indexes year. $\ln(yield_{it})$ is the natural logarithm of a crop yield. $\ln(ozone_{it})$ is the natural logarithm of an ozone measure. x_{kit} are control variables including temperature variables and total rainfall and their squared terms to take into account of the nonlinear impact of climate variables on crop yield (Schlenker and Roberts, 2009). Year dummy variables are also included to account for year fixed effects. θ_i is a set of county fixed effects which control for time-invariant heterogeneity such as farm practice and soil quality, ϵ_{it} is an error term assumed to be normally distributed with zero mean. The coefficient of interest, β_1, is interpreted as elasticity of crop yield with respect to ozone.

The identification assumption for equation (16.1) is that $E[ozone_{it}\epsilon_{it}|\theta_i, x_{it}] = 0$. In other words, after controlling for county fixed effects, temperature and rainfall, ozone measure and an error term should be uncorrelated. This assumption might not hold, however, if farmers respond to the negative impacts on crop yield from ozone by, for example, applying more fertilizer to increase yield. Omitting fertilizer variable from equation (16.1) would cause β_1 to bias toward zero.

I address the omitted variable bias by using instrumental variable and two-stage least square (2SLS) approach. In 2003 or 2004, 22 states[1] in the eastern U.S. must be in compliance with the NOx Budget Trading Program (NBP) created to reduce emissions of NOx from power plants in the eastern U.S. during the summer months.

[1]Eight states including Connecticut, Delaware, Maryland, Massachusetts, New Jersey, New York, Pennsylvania, and Rhode Island, as well as the District of Columbia started their compliance on May 1, 2003. Alabama, Illinois, Indiana, Kentucky, Michigan, North Carolina, Ohio, South Carolina, Tennessee, Virginia, and West Virginia began compliance on May 31, 2004

Under this program, an emission cap is set and tradable allowances are allocated to each affected source. EPA (2008) reported that ozone season NOx emissions have decreased by more than 60% in 2008 after the program is implemented. The instrumental variable is a dummy variable indicating years before and after the NBP program[2]. For a county within the state in compliance with the program, the dummy variable equals 1 from the year of compliance on and equals 0 for the years before the compliance. Two assumptions are required for the instrument to be valid. First, the instrument must be correlated with ozone measures. I will test this assumption in the first-stage regression. Second, the instrument should not correlate with the omitted variable affecting crop yields such as farmer's behavior. This assumption is likely to be valid because the variation of the county's compliance year should not be correlated with farmer's response to negative effect of ozone.

[2]A caveat for using this instrument is that the program might not be effective in reducing ozone in the downwind states if the upwind states do not in compliance with the program. For the future work, wind direction data and NOx emission from coal-fired power plants data can be used to construct an alternative instrument. Since wind direction is exogeneous and NOx is the key ingredient in ozone formation, NOx concentration from upwind states can be used as a instrument for ozone concentration in the downwind states.

Chapter 17

Results

Table 17.1 reports results for soybean and corn yield regressions using a mean ozone as an ozone measure. I allow for disturbance terms ϵ_{it} to be correlated within the same county and year. The standard errors reported in the estimation results are therefore a robust variance estimator that is clustered by each county-year combination. All of the variables have expected signs and are significant at the 1% level. Columns (1) and (4) show the estimated effect of mean ozone on soybean and corn yield respectively without controlling for any covariate but time and county fixed effects. Ozone has a negative impact on crop yield. The estimated elasticity of yield with respect to mean ozone is -0.55 for soybean and -0.64 for corn. In other words, a one percent increase in mean ozone decreases soybean and corn yield by 0.55 and 0.64 percent respectively. These estimated effects, however, could be overestimated because high temperature and low rainfall would have a negative impact on crop yield as well. I control for these variables in column (2) and (5). As expected, the elasticites of crop yield decrease significantly to -0.44 for soybeans and to -0.49 for corn. The coefficients of temperature variable and its squared term indicate the non-linear impacts of temperature on yield. Consistent with the estimates from Schlenker and Roberts (2009), yields increase with temperature up to 83 °F for corn, 88 °F for soybeans. Rainfall has positive effects on both yields.

As mentioned earlier, I use a dummy variable indicating years before and after the NBP as an instrument for ozone measures to address the omitted variable bias problem. Table 17.2 shows the first-stage regressions for both crops. The mean ozone measure is regressed with the instrumental variable and all covariates. The estimated coefficients of the instrument for both crops are negative and highly significant indicating that mean ozone is negatively correlated with the instrument. Column (3) and (6) show the estimated results from instrumental variable regressions. As expected, the estimated coefficients are bigger than those estimated from fixed effect model. This implies a downward bias by omitted variables. The estimated elasticity of yield with respect to mean ozone is -0.60 for soybean and -0.57 for corn. Consistent with the literature, soybeans are more sensitive to yield loss caused by ozone than corn.

However, my estimated elasticities are higher than what was found in Westenbarger and Frisvold (1995) where the estimated elasticity is about -0.54 for soybeans and -0.19 for corn.

To put these estimated elasticities into perspective, I calculate back-of-the-envelope value of crop losses as a result of an increase in one standard deviation of the mean ozone measure from its mean. This number is equivalent to 11% [1]. Using the value of soybean and corn production in 2008 which are about 30 and 50 billion dollars[2], one standard deviation increases in mean ozone would damage soybean and corn by 1.96 and 3.10 billion dollars, respectively.

Instead of using the mean ozone measure, table 17.3 and table 17.4 report the regression results using a maximum ozone and a maximum 8-hour ozone measure respectively. From both tables, a similar pattern emerges. All of the variables have expected signs and are significant at the 1% level. The estimated elasticities are lower when control for covariates in column (2) and (5). The instrumental variable approach increases the estimated elasticities in column (3) and (6). From the instrumental approach, the estimated elasticties of yield with respect to maximum and maximum-8hr ozone measure are similar. The estimated elasticties are -0.73 and -.64 for soybeans and corn respectively. These estimates can be used for the back of the envelop calculation of the benefits from NBP. Butler and Lee (2007) estimated that in 2007 the NBP reduces 8-hour ozone concentration on average by 10% below the level seen before the program in 2003. Using value of soybean and corn production in year 2007, the NBP reduces the value of crop losses by 2.19 billion dollars for soybeans and 3.2 billion dollars for corn.

---

[1]From the dataset, the mean and the standard deviation of the mean ozone measure for both crops are 0.046 ppm and 0.005 ppm respectively. The one standard deviation from its mean is $0.0050.046 \times 100 \approx 11\%$.

[2]http://www.nass.usda.gov/QuickStats/indexbysubject.jsp?Passname=&Passgroup=Crops +%26+PlantsPasssubgroup=Field+Crops#top

Table 17.1: Estimated effect of mean ozone on crop yields.

	Dependent variable					
	Log of soybean yields			Log of corn yields		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	FE	**FE**	**IV**	**FE**	**FE**	**IV**
Log of weighted mean ozone (ppm)	-0.5561***	-0.4368***	-0.6052***	-0.6421***	-0.4988***	-0.5706***
	(0.0804)	(0.0807)	(0.1780)	(0.0866)	(0.0823)	(0.1611)
Mean temperature (°F)		0.2630***	0.2814***		0.1262***	0.1443***
		(0.0393)	(0.0394)		(0.0615)	(0.0303)
Mean temperature^2 (°F)		-0.0015***	-0.0016***		-0.0007***	-0.0008***
		(0.0003)	(0.0003)		(0.0003)	(0.0002)
Rainfall (mm)		0.0001***	0.0001***		0.0001**	0.0001***
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Rainfall^2(mm)		-0.0000	-0.0000**		-0.0000	-0.0000**
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Constant	1.7373***	5.0681***	8.0323***	2.6117***	2.0835	6.0276**
	(0.2504)	(1.3623)	(2.2728)	(0.2745)	(2.1102)	(2.4536)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,495	1,495	1,495	1,919	1,919	1,919
No. of counties	102	102	102	132	132	132
R-squared	0.214	0.277	n.a.	0.307	0.338	n.a

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

Table 17.2: First-stage regression for instrumental variable regression of impact of mean ozone on crop yield

	First-stage regression	
	Soybeans	**Corn**
	(1)	**(2)**
Instrument variable	-0.269***	-0.1459***
	(0.0086)	(0.0082)
Mean temperature (°F)	0.0268**	0.0179**
	(0.0127)	(0.0090)
Mean temperature^2 (°F)	-0.0000	-0.0000
	(0.0000)	(0.0000)
Rainfall (mm)	0.0000	0.0000
	(0.0000)	(0.0000)
Rainfall^2 (mm)	0.0000**	0.0000**
	(0.0000)	(0.0000)
Constant	0.4266***	0.3912***
	(0.0421)	(0.0296)
Year fixed effect	Yes	Yes
Obs.	1,459	1,919
No. of counties	102	132

*** $p<0.01$, ** $p<0.05$
Robust standard errors in parentheses

Table 17.3: Estimated effect of maximum ozone on crop yields.

	Dependent variable					
	Log of soybean yields			Log of corn yields		
	(1) FE	(2) FE	(3) IV	(4) FE	(5) FE	(6) IV
Log of maximum ozone (ppm)	-0.5567***	-0.4356***	-0.7303***	-0.6317***	-0.4615***	-0.6458***
	(0.0880)	(0.0913)	(0.2039)	(0.0920)	(0.0909)	(0.1834)
Mean temperature (°F)		0.2332***	0.2678***		0.0986**	0.1232***
		(0.0396)	(0.0398)		(0.0416)	(0.0302)
Mean temperature^2 (°F)		-0.0015***	-0.0016***		-0.0006**	-0.0007***
		(0.0003)	(0.0003)		(0.0002)	(0.0003)
Rainfall (mm)		0.0001***	0.0001***		0.0001**	0.0001**
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Rainfall^2 (mm)		-0.0000	-0.0000*		-0.0000	-0.0000*
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Constant	1.8587***	4.9915***	10.2805***	2.7884***	0.3885	6.0805**
	(0.2552)	(1.3787)	(2.3361)	(0.2713)	(2.1193)	(2.5079)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,495	1,495	1,495	1,919	1,919	1,919
No. of counties	102	102	102	132	132	132
R-squared	0.208	0.273	n.a.	0.299	0.331	n.a

*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses

Table 17.4: Estimated effect of maximum 8-hrs ozone on crop yields.

	Dependent variable					
	Log of soybean yields			Log of corn yields		
	(1) FE	(2) FE	(3) IV	(4) FE	(5) FE	(6) IV
Log of maximum 8-hrs ozone (ppm)	-0.5700***	-0.4473***	-0.7259***	-0.6450***	-0.4878***	-0.6237***
	(0.0846)	(0.0860)	(0.2037)	(0.0889)	(0.0856)	(0.1775)
Mean temperature (°F		0.2226***	0.2642***		0.1143***	0.1250***
		(0.0393)	(0.0396)		(0.0213)	(0.0301)
Mean temperature^2 (°F)		-0.0015***	-0.0016***		-0.0006***	-0.0007***
		(0.0003)	(0.0003)		(0.0002)	(0.0002)
Rainfall (mm)		0.0001***	0.0001***		0.0001**	0.0001***
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Rainfall^2 (mm)		-0.0000	-0.0000**		-0.0000	-0.0000**
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Constant	1.7580***	5.0420***	10.2731***	2.6774***	4.2408**	6.0655**
	(0.2548)	(1.3681)	(2.3427)	(0.2717)	(2.0988)	(2.4900)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,495	1,495	1,495	1,919	1,919	1,919
No. of counties	102	102	102	132	132	132
R-squared	0.211	0.275	n.a.	0.303	0.334	n.a

*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses

Chapter 18

Conclusion

Ground-level ozone is harmful not only to human health but also harmful to crops. This paper examines the impact of ozone on two of the U.S. major crops, soybeans and corn. Using ground-level ozone concentration data from the EPA's monitoring network, county-level crop yield data from USDA and instrumental variable approach, I find that ozone substantially reduces crop yields. In particular, the elasticity of crop yield with respect to various ozone measures are ranging from -0.60 to -0.73 and -0.57 to -0.64 for soybeans and corn respectively. These estimats are a bit bigger than previously found in the comparable studies.

Using the estimated elasticities, a back of the envelope calculation indicates that the value of crop production losses for corn and soybeans due to a standard deviation increase of mean ozone is about $5 billion dollars in 2008. Using Butler and Lee (2007) estimates of reduction of ozone by NOx Budget Trading Program during 2003-2007, my estimated elasticity of crop yield with respect to maximum 8-hour ozone concentration translate into reduction of corn and soybean losses by about 2.19 billion and 3.2 billion respectively.

Part IV

References

References for Part I

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

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

California Energy Commission (2005). *Integrated Energy Policy Report.* Sacramento, California.

Cline, W. R. (1992). *The economics of global warming.* Washington: Institute for International Economics.

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.

Deschênes, O., and M. Greenstone. (2007). "Climate Change, Mortality and Adaptation: Evidence from annual fluctuations in weather in the U.S." MIT Department of Economics Working Paper No. 07-19.

Energy Information Administration (EIA). (2008). *State Energy Data System.* Washington, D.C.

Fisher, F. M., and C. Kaysen. (1962). *A Study in Econometrics: The Demand for Electricity in the US.* North Holland Publishing Company.

Franco, G., and A. Sanstad. (2008). "Climate change and electricity demand in California." *Climatic Change* 87:139-151.

Hanemann, W. (1984). "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52: 541-561.

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

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

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.

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.

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.

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.

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

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.

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

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.

References for Part II

ACEEE (2008). *The 2008 State Energy Efficiency Scorecard.* ACEEE Report E086.

Appliance Standards Awareness Project (2006). www.standardsasap.org

Arellano, M., and S.R. Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Reviews of Economics Studies* 58, 277-298.

Baltagi, B.H. and Griffin, J.M. (1997). "Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline." *Journal of Econometrics* 77, 303-327.

Baltagi, Badi H. (2001). *Econometric analysis of panel data.* 2nd edition: John Wiley & Sons, Ltd

Baltagi, B.H., Bresson, G. and Pirotte, A. (2002). "Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: Some empirical evidence from US electricity and natural-gas consumption." *Economic Letters* 76, 375-382.

Berndt, Ernst R. (1996). *The Practice of Econometrics: Classic and Contemporary.* : Addison Wesley.

Bion, Howard and W. Prindle. (1991). *Better Building Codes Through Energy Efficiency.* Washington, D.C: Alliance to Save Energy.

Brookes, Len (1990)."The Greenhouse Effect: The Fallacies in the Energy Efficiency Solution." *Energy Policy* Vol.18 No. 2, 199-201

Building Codes Assistance Project (BCAP) (2009). Code Status and maps. *http://bcap-energy.org/node/5*

Building Codes Assistance Project. (2008). *Residential Building Energy Codes - Enforcement & Compliance Study.* Prepared for and funded by the North American Insulation Manufacturers Association, October.

Bureau of Economic Analysis (2009). State annual tables, Washington D.C.: Department of commerce.

California Energy Commission (2007). "California Energy Demand 2008-2018" *CEC Report* 200-2007-015-SF2.

Dahl, Carol. (1993). *A Survey of Energy Demand Elasticities in Support of the Development of the NEMS.* Washington, DC: U.S. Department of Energy.

Energy Information Administration (2008). "Emissions of Greenhouse Gases Report." EIA Report #: DOE/EIA-0573

Energy Information Administration (2009). State Energy Data System *http://www.eia.doe.gov/emeu/states/$_s$eds.html.*

Geller, Howard (1995). *National Appliance Efficiency Standards: Cost-Effective Federal Regulations.* American Council for an Energy-Efficiency Economy, Washington D.C.: Report A951

Gillingham, Kenneth, R. G. Newell, and K. L. Palmer. (2006). "Energy Efficiency Policies: A Retrospective Examination." *Annual Review of Environment and Resources* 31, November.

Horowitz, Marvin J. (2007). "Changes in electricity demand in the United States from the 1970 s to 2003"'. *The Energy Journal* 28(3): 93-119.

Horowitz, Marvin J. (2004). "Electricity Intensity in the Commercial Sector: Market and Public Program Effects." *The Energy Journal* Vol 25 No. 2, 115-137.

Houthakker H. S. (1962). "Electricity Tariffs in Theory and Practice." Electricity in the United States. Amsterdam: North Holland Publishing Co.

Houthakker, H.S., P.K. Verleger, and D. Sheehan (1974). "Dynamic demand for gasoline and residential electricity." *American Journal of Agricultural Economics* 56, 412 -418.

Hsiao, C. (1986). A*nalysis of panel data.* Cambridge: Cambridge University Press.

Jones, Ted, D. Norland, and W. Prindle.(1998). *Opportunity Lost: Better Energy Codes For Affordable Housing and a Cleaner Environment.* Alliance to Save Energy, Washingto D.C.

Gillingham, Kenneth, R. Newell, and K. Palmer (2004). *Retrospective Examination of Demand-Side Energy Efficiency Policies.* Resources for the Future, Washington D.C.

Hanemann, W.M. (1984). "Discrete/continuous models of consumer demand."' *Econometrica*: pp. 541-561.

Khazzoom, D.J. (1980). "Economic Implications of Mandated Efficiency Standards for Household Appliances." *The Energy Journal* Vol. 1 No.4, 21-39.

Loughran, David S. and Jo Kulick (2004). "Demand-side Management and Energy Efficiency in the United States." *The Energy Journal* Vol. 25, No. 1, 19-42.

Maddala, G.S., Trost, R.P., Li, H., and Joutz, F. (1997). "Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators." *Journal of Business and Economic Statistics* 15, 90-100.

Marshall, Lynn, and Tom Gorin, Principal Authors. (2007). "California Energy Demand 2008-2018 - Staff Revised Forecast."' California Energy Commission Staff Final Report # CEC-200-2007-015-SF2, November.

Mosenthal Philip and Jeffrey Loiter. (2007). "National Action Plan for Energy Efficiency. Guide for Conducting Energy Efficiency Potential Studies." Prepared by Optimal Energy, Inc. Website: www.epa.gov/eeactionplan

Nadel, Steven, A. deLaski, M. Eldridge and J. Kleisch (2006). "Leading the Way: Continued Opportunities for New State Appliance and Equipment Efficiency Standards. American Council for an Energy-Efficient Economy and Appliance Standards Awareness Project." Washigton, D.C.:Report Number ASAP-6/ACEEE-A062.

Nadel, Steven, A. deLaski, M. Eldridge, and J. Kleish (2005). "Leading the Way: Continued Opportunities for New State Appliance and Equipment Efficiency Standards. American Council for an Energy-Efficiency Economy and Appliance Standards Awareness Project."' Washington D.C.: Report A051

NOAA (2005). "State heating and cooling degree days 1960-2000." Private communication to David Rich.

Prindle, William, N. Dietsch, R. Neal Elliott, M. Kushler, T. Langer, and S. Nadel (2003). "Energy Efficiency's Next Generation Innovation at the State Level. American Council for an Energy-Efficient Economy." Washington, D.C.: Report No. E031

Sanstad, Alan H. (2007). "The Evaluation of Residential Utility Demand-Side Management Programs in California." In Richard D. Morgenstern and William A. Pizer, Eds., Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe, and Japan. Washington, DC: Resources for the Future.

Sanstad, Alan H., and W. Michael Hanemann, Maximilian Auffhammer. (2006). "End-Use Energy Efficiency in a 'Post-Carbon' California Economy: Policy Issues and Research Frontiers." In Alexander E. Farrell and W. Michael Hanemann, Eds., Managing Greenhouse Gas Emissions in California. San Francisco, California: The Energy Foundation.

Saunders, H.(1992)."The Khazzom-Brookes Postulate and Neo-classical Growth." *The Energy Journal* Vol. 13 No. 4, 131-148.

Sudarshan, A. and Sweeney, J. (2008). Deconstructing the “Rosenfeld Curve”. Stanford PIEE Working Paper Series.

United States Census Bureau (2009). Housing Units Authorized by Building Permits. Website: http://www.census.gov/const/www/C40/table2.html

Wooldridge, Jeffrey M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: The MIT Press.

References for Part III

Auffhammer, M. and R. Kellogg. (2009). “Clearing the air? the eects of gasoline content regulation on air quality.” Center for the study of energy markets working paper.

Bernardo F., T. Rudor, C. L. Mulchi, E. H. Lee, R. Rowland, and R. Pausch. (1996). “Effects of enhanced o3 and co2 enrichment on plant characteristics in wheat and corn. *Environmental Pollution.* 94(1):53-60.

David A. Westenbarger and George B. Frisvold. (1995) “Air pollution and farm-level crop yields: An empirical analysis of corn and soybean.” *Agricultural and Resource Economic Review.* 24:156-164

Environment Protection Agency. (2006).“Air quality criteria for ozone and related photochemical oxidants.” Technical report. Office of Research and Development :Washington,DC.

Environment Protection Agency. (2006). “Ozone and your health EPA.” Technical report. Washington, DC.

Ficus, E., C.D. Reid, J.E. Miller, and A. Heagle. (1997) “Elevated CO2 reduces O3 flux and O3-induced yield losses in soybeans: possible implications for elevated CO2 studies.” *Journal of Experimentl Botany.* 48:307-313

Fiscus, E., F.L. Booker, and K.O. Bureky.(2005) “Crop responses to ozone: uptake, modes of action, carbon assimilation and partitioning.” *Plan Cell Environ.* 28: 997-1011

Garcia, P., B. L. Dixon, J. W. Mjelde, and R. M. Adams.(1986). “Measuring the benefits of environmental change using a duality approach: The case of ozone and Illinois cash grain farms.”’ *Journal of Environmental Economics and Management.* 13(1): 69-80

Heagle,A.(1989). “Ozone and crop yield.” *Annual Review of Phytopathology.*

27:397-423.

Krupa, Sagar (1997) " Air Pollution, People, and Plants." *American Phytopathological Society.*

Lefohn, A. S.and V.C. Runeckles.(1967) "Establishing standards to protect vegetation ozone exposure/dose considerations." *Atmospheric Environment.* 21(3):561

Leung, S.K. , W. Reed, and S. Geng. (1982). "Estimations of ozone damage to selected crops grown in southern california." *Journal of the Air Pollution Control Association.* 32:160-164.

Likens, G., M. Rury, T. Butler, F. Vermeylen, and B. Lee. (2007). Regional impacts of reduced nox emissions on ozone concentrations in the eastern usa. Technical report: EPA Clean Air Markets Division.

Mackee, I.F, Bullimore J.F., and Long S.P. (1997). "Will elevated co2 protect the yield of wheat from 03 damage?" *Plant, Cell and Environment.* 20:77-84.

Maggs, R. and M.R. Ashmore.(1998). "Growth and yield responses of pakistan rice cultivars to O3 and NOx."' *Environmental Pollution.* 103:159-170.

Mauzerall, Denise L. and Xiaoping Wang. (2001). "Protecting agricultural crops from the eects of tropospheric ozone exposure: Reconciling science and standard setting in the united states, europe, and asia." *Annual Review of Energy and the Environment.* 26:237-268.//

Wahid, A. , R. Maggs, S. R. A. Shamsi, J. N. B. Bell, and M. R. Ashmore. (1995). "Effects of air pollution on rice yield in the pakistan punjab." *Environmental Pollution.* 90(3):323-329.

Westenbarger, A. D. and G. B. Frisvold. (1994). Agricultural exposure to ozone and acid precipitation. Atmospheric Environment, 28:2895-2907.

Westenbarger, A. D. and G. B. Frisvold. Air pollution and farm-level crop yields: An empirical analysis of corn and soybean. (1995). *Agricultural and Resource Economic Review.* 24:156-164.

Robert D. Rowe and Lauraine G. Chestnut (1985) "Economic Assessment of the effects of Air Pollution on Agricutural Crops in the San Joaquin Valley." *Journal of the Air Pollution Control Association* (35) : 728-734.

Schlenker, Wolfram and Michael J. 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): 15594-15598.

Supriya, T., M. Agrawal, and W. J. Manning.(2005). "Assessing the impact of ambient ozone on growth and productivity of two cultivars of wheat in india using three rates of application of ethylenediurea (edu)." Environmental Pollution. 138(1):153-160.

Zlatev, Z., I. Dimov, Tz. Ostromsky, G. Geernaert, I. Tzvetanov, and A. Bastrup-Birk. (2001). "Calculating losses of crops in denmark caused by high ozone levels." *Environmental Modeling and Assessment.* 6(1):35-55.