Essays on Water and Energy Economics

Skylar Olsen

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Hendrik G. Wolff, Chair

Robert Halvorsen

Seik Kim

Joseph H. Cook

Program Authorized to Offer Degree:
Department of Economics

University of Washington

Abstract

Essays on Water and Energy Economics

Skylar Olsen

Chair of the Supervisory Committee:
Assistant Professor Hendrik G. Wolff
Department of Economics

The preferences and behavior of economic agents is often highly nuanced and heterogenous, having large impacts on the costs and benefits of different methods employed to acheive environmental goals, such as the production of green energy and the efficient allocation of water to residential use. The first chapter of this dissertaion studies local housing and labor market effects around US nuclear power plants from 1970 to 2000 through three major events: initial construction of the power plant, the Chernobyl accident, and plant closings. Heavily effected by preference-based sorting and mirgration, nuclear power plant placement lowers unemployment, and increases the education and skill level within the local labor pool. However, nuclear power plants also pose negative externalities: conditional upon labor market improvements, housing values around 10 miles of the nuclear plants decrease with plant placement and increase with plant closures. The 1986 Chernobyl nuclear accident in the Ukraine serves as an information shock to US residents. Intercontinental spillover effects result in lower US home ownership rates around US plants and reductions in housing values in the plume exposure pathway. The final chapter examines the impact of non-pecuniary incentives stemming from the behavioral economics literature on water demand. In a randomized field experiment social comparisons are found to significantly decrease water demand with substantial heterogeneity both across and within utilities. Higher users are more responsive to the program and there are important interactions between social norms

and existing utility conservation programs. Water resources and energy supply face stress due to population growth, rising incomes, and climate change and these stressors will only increase in the future. This dissertation addresses key issues: influencing the demand for water and the side effects of commercial nuclear power supply and aims to increase knowledge and aid the public policy of both water management and the supply of green energy.

TABLE OF CONTENTS

Page

List of Tables iii

List of Figures v

Chapter 1: Nuclear Reactors in the US: Housing Values, Sorting, Migration, and Employment 1

1.1 Introduction 2

1.2 Previous Literature 7

1.3 Data 12

1.4 Treatment Group 13

1.4.1 Plant Placement 13

1.4.2 Plant Closures 15

1.4.3 The Chernobyl Accident 17

1.4.4 Defining Treatment 17

1.5 Control Group 23

1.6 Net changes in labor market and demographic mix 40

1.6.1 Methodology 40

1.6.2 Results: Plant placement 41

1.6.3 Results: Plant closures 44

1.6.4 Results: Chernobyl accident 46

1.7 Amenity value of nuclear power plants 49

1.7.1 Methodology 49

1.7.2 Results: Plant placement 50

1.7.3 Results: Plant closures 52

1.7.4 Results: Chernobyl accident 53

1.8 Conclusion 53

Chapter 2: Heterogeneous Responses to Social Norms for Water Conservation 54

2.1 Introduction 55

2.2 Theoretical Background 58

2.3 Existing Literature . . . 60
2.4 Background & Data . . . 63
2.4.1 WaterSmart Software . . . 63
2.4.2 Data . . . 66
2.4.3 Utilities . . . 67
2.4.4 Experimental Design . . . 70
2.5 Estimation Strategy . . . 74
2.5.1 Average Treatment Effects . . . 74
2.5.2 Durability . . . 76
2.5.3 Treatment Effect Heterogeneity . . . 82
2.5.4 Interactions with Existing Conservation Programs . . . 89
2.6 Future Work . . . 94
2.7 Conclusion . . . 95

Bibliography . . . 98

Appendix A: Appendix for Chapter 1 . . . 106
A.1 Challenges to US Nuclear Energy and the Federal Strategy . . . 107
A.2 Surveys and Stated Choice Studies: the Perception of Nuclear Power . . . 109
A.3 Discussion of Hedonic Method in the Current Context . . . 111
A.4 Identification . . . 113
A.4.1 Defining Treatment and Geographic Sample Exclusions . . . 113
A.4.2 Endogeneity due to Placement . . . 114

Appendix B: Appendix for Chapter 2 . . . 119
B.1 Pilots . . . 120
B.2 Irregular Period Lengths and Staggered Reads . . . 120
B.3 Matching Weather Data to Irregular Meter Read Periods . . . 126
B.4 Cleaning other outside data sources . . . 131
B.5 Water Use . . . 133
B.5.1 Distribution and Outliers . . . 133
B.5.2 Balance in Baseline Water Use Between Treatment and Control . . . 135
B.5.3 Water Use Time Series . . . 140
B.6 Conditional Average Treatment Effects . . . 143

LIST OF TABLES

Table Number		Page
1.1	Treatment group sample sizes by nuclear power plant set	22
1.2	Sample means of the treatment and baseline control groups for each of the nuclear power plant sets used to evaluate the effect of placement, closures, or the Chernobyl accident.	26
1.3	Summary statistics for plant placement set by 10-mile treatment group and control subsamples.	30
1.4	Summary statistics for successful plant placement set by 10-mile treatment group and control subsamples.	31
1.5	Summary statistics for plant closure set by 10-mile treatment group and control subsamples.	32
1.6	Summary statistics for Chernobyl accident set by 10-mile treatment group and control subsamples.	33
1.7	Summary statistics for plant placement set by 10-mile treatment group and weighted control subsamples.	36
1.8	Summary statistics for succesfull plant placement set by 10-mile treatment group and weighted control subsamples.	37
1.9	Summary statistics for plant closure set by 10-mile treatment group and weighted control subsamples.	38
1.10	Summary statistics for Chernobyl accident set by 10-mile treatment group and weighted control subsamples.	39
1.11	Net effect of nuclear power plant placement	42
1.12	Net effect of successful nuclear power plant placement	45
1.13	Net effect of nuclear power plant closures	47
1.14	Net effect of the Chernobyl accident	48
1.15	The amenity/disamenity value of successful plant placement	51
1.16	The amenity/disamenity value of plant closures	52
2.1	Summary Statistics By Pilot	68
2.2	Utility Conservation Programs	69
2.3	Differences in Means Between Control and Treatment Groups	72
2.4	Sample Sizes	72
2.5	Motivating the Rank Preservation Assumption	74

2.6 Average Treatment Effects . . . 75
2.7 Average Treatment Effects: Exclude first period . . . 79
2.8 Heterogeneity: Baseline Water Use . . . 87
2.9 Heterogeneity: Ideology . . . 88
2.10 Heterogeneity: Housing Values . . . 89
2.11 Logit Regressions for Conservation Programs - Watertown . . . 91
2.12 Interactions with Conservation Programs - Watertown . . . 93

B.1 Initial Sample Sizes . . . 120
B.2 Period Length: high and low quantiles by pilot . . . 121
B.3 Short and long read periods relative to pilot . . . 123
B.4 Observation counts by month and pilot. . . . 123
B.5 Observation counts by reading period and pilot. . . . 125
B.6 Ensuring a balanced panel: identify multiple obs within year-period. . . . 126
B.7 Percent of Water Read Observations Missing Corresponding Weather Data . 127
B.8 Summary Statistics By Pilot: Baseline water use and monthly weather . . . 127
B.9 Percent of Non-missing Home Feature Data Provided by WaterSmart . . . 131
B.10 Summary Statistics By Pilot: Baseline water use and HH characteristics . . 133
B.11 Percent of Households Missing Data . . . 133
B.12 Water Use: High and low quantiles by pilot . . . 135
B.13 Median Water Use by Season . . . 135
B.14 Outliers in High Water Use (observation counts) . . . 136
B.15 Outliers in Low Water Use (observation counts) . . . 136
B.16 Differences in Means Between Control and Treatment Groups . . . 137
B.17 Differences in Means Between Control and Treatment Groups for Hydroburg after Matching . . . 138
B.18 Heterogeneity: Baseline Water Use . . . 144
B.19 Heterogeneity: Ideology . . . 145
B.20 Heterogeneity: Housing Values . . . 146

LIST OF FIGURES

Figure Number Page

1.1 Nuclear Plants Placed After 1971 and Before 1980 15
1.2 Nuclear Plants Placed After 1971 and Before 1980, which Successfully Operated 16
1.3 Nuclear Plants Closed in the 1990s 16
1.4 Nuclear Plants Operating Before 1980 and Through 1990 18
1.5 Nuclear Plants Placed After 1971 and Before 1980 19
1.6 Nuclear Plants Placed After 1971 and Before 1980, which Successfully Operated 20
1.7 Nuclear Plants Closed in the 1990s 20
1.8 Nuclear Plants Operating Before 1980 and Through 1990 21
1.9 Defining baseline control tracts: Successful placement set 24
1.10 Distributions of Population Density and Vancancy Rate for Whole Sample and Treatment Group of Plant Placement Set 25
1.11 The Min/Max Subsample: Successful placement set 28
1.12 The Min/Max Weighted Subsample: Successful placement set 35

2.1 Home Water Report 64
2.2 Online Web Portal 65
2.3 Baseline Water Use by Pilot 71
2.4 Durability of Treatment Effects - Percentage 78
2.5 Durability of Treatment Effects - Gallons per day 81
2.6 Quantile Treatment Effects 85

B.1 Period Length Distributions: Identifying short and long periods 122
B.2 Frequency of water reads by pilot and day of the month 124
B.3 Evapotranspiration 128
B.4 Average Maximum Daily Temperature 129
B.5 Monthly Precipitation 130
B.6 Distributions of gallons per day by season 134
B.7 Quantile-quantile plots to explore effect of matching algorithm on balance . . 139
B.8 Water Use by Pilot 141
B.9 Year-Over-Year Change in Water Use by Pilot 142

ACKNOWLEDGMENTS

The journey through graduate school to the eventual procurement of my PhD provided some of the most rewarding and the most trying experiences of my life. The early years were composed of rigorous coursework, testing one's abilities to reason and organize one's thoughts, to solve multifaceted problems with pages and pages of mathematics, and to diligently seek answers in textbooks, peers, and professors. The dissertation years were laden with doubt, searching, hours coding and pouring over results, and no small amount of stress.

I would have never reached the end of this road without valuable collaboration with peers, most specifically my co-author and good friend Danny Brent, and the advice of my committee: professors Seik Kim, Robert Halvorsen, Joe Cook, and most importantly my chair and advisor Hendrik Wolff.

I owe my continued sanity to my wonderful family and friends: to my parents whose unending faith in me combat moments of doubt, who helped me re-find purpose and outside options when I found myself on a path I no longer wanted to walk; to the analytics team at Zillow who became that purpose, to friends like Danny Brent and Joe Dieleman (as well as many others) who laughed and vented with me along the way, and finally, and most importantly, to the love and partner of my life, Zachary Ohnemus, for unflagging dedication, for his shoulder, his calm and confident bearing, for his humor, and for taking care of me through the best and worst of times.

DEDICATION

to my heart, Zachary

Chapter 1

NUCLEAR REACTORS IN THE US: HOUSING VALUES, SORTING, MIGRATION, AND EMPLOYMENT

1.1 Introduction

The cost-benefit analysis of place-based investments and long-term localized interventions to surrounding communities continues to be a popular topic in the environmental economics literature. Papers in the last few years have focused on the large scale development of wind power (Heintzelman and Tuttle, 2012); the local impact of polluting operations such as coal-fired plants or concentrated feeding operations (Davis, 2011; Kuethe and Keeney, 2012), and the location of landfills or other waste sites (Braden et al., 2011), to list only a few topics in this diverse set of literature. While making significant contributions, there remains a need for deeper understanding of the influence of local preferences and opposition to the initial site choice and the long-term impacts of place-based long-term interventions. Importantly, these long-term impacts include the preference- and skill-based sorting of individuals away from or towards the site. This non-random siting and post-placement sorting can mitigate local negative externalities and materially influence the cost-benefits analysis of these interventions. Nuclear power plants, as controversial long-term investments in predominantly rural areas where the plant constitutes a significant employer, offer an ideal situation to study this long-term sorting and simultaneous changes in local housing and labor markets.

Currently the world is experiencing the greatest change in the commercial nuclear power landscape seen in decades. In March 2011, the Fukushima accident generated massive concerns about nuclear energy. Since then, Japans commercial nuclear energy production dwindled to practically zero. In Germany, nine of seventeen plants were immediately shutdown and a commitment was made to close the remainder within a decade. The first United States closures since 1997, three plants permanently shut down during the first half of 2013 and another is expected to close in the fourth quarter of 2014.

In the US, despite these closures, nuclear power plays an important part in the discussion around US energy policy goals.[1] Five nuclear reactors are currently under construction and

[1]See Appendix A for a discussion of the challenges facing US commercial nuclear power and the current strategy of the US federal government to increase nuclear energy capacity.

28 license applications are under review by the US Nuclear Regulatory Commission (NRC). The first nuclear power reactor to be connected to the US grid since 1996 is expected to be ready for commercial energy production in December 2015. The US is not alone in revamping the nuclear arena. Globally, 62 nuclear reactors are currently under construction, 167 are in the planning stage, and 317 are proposed. These plans, however, experience massive opposition by local communities and as a result the siting of potential nuclear plants is an extremely difficult task. This paper will help to better understand the situation of local communities by using revealed preferences.[2]

We study the period from 1970 through 2000, which parallels the current evolving climate surrounding nuclear power: rapid growth fueled by strong governmental support in the 1970s was eventually curtailed, due to plant location controversies and fear resulting from the Chernobyl accident, and followed by the closures of select plants. This presents the first paper to estimate both the economic and the demographic impact of three major events (treatments) in the lives of nuclear power plants: the initial placement of the power plant, a global nuclear accident (we use the Chernobyl accident as an information shock), and plant closures. Previous hedonic studies analyzed individual plants and have yielded idiosyncratic results on property values (Folland and Hough, 2000; Clark and Allison, 1999; Clark et al., 1997; Gamble and Downing, 1982; Nelson, 1981). Except for Folland and Hough (2000), who focus on the value of agricultural land, none of the above studies have looked at new plant openings or plant closings. Rather these studies either compare property values in a cross section or use other types of treatments, such as the transportation of nuclear waste, to study the effect of nuclear plants on home values. Cross sectional analysis, we later argue, can be problematic due to sorting. In the case of using other types of treatments like waste transport, these treatments are minor besides plant placements and closures.

We construct the most spatially comprehensive dataset tracking the history of US nuclear power plants from the construction phase, through the generation of electricity, and in some

[2]See Appendix B for a discussion of stated preference techniques applied to illicit opinions about nuclear energy.

cases plant closures. We recorded the longitude and latitude of each nuclear power plant as well as their potential plume exposure pathway in the case of an accident. We spatially merge this plant level dataset to tract level demographic, labor market, and housing market variables from the US Census Neighborhood Change Database[3]. This allows us to estimate endogenous sorting by analyzing such features as migration decisions and changes in the demographic makeup or the occupational structure of the local labor market. We show that a nuclear power plant presents a large local employer providing many high skill and high wage earning jobs. As power plants are typically located in rural areas, plant siting induces strong simultaneous changes in the local labor market and the demographic mix. Finally, we quantify a nonmarket amenity value of nuclear plants, by carefully estimating the local housing value discount or premium associated with plants conditional on our sorting measures.

We differentiate (a) the net effect of a nuclear power plant on housing and other outcomes from (b) this pure nonmarket amenity effect.[4] These estimated values can be very different. Whereas house values could increase around nuclear power plants due to the inflow of highly paid technicians operating the facility, or decrease due to a loss of overall population and demand, the pure amenity effect of the power plant is different from this net effect. A priori the sign on this amenity effect is not determined. The amenity value could be positive or negative depending how residents value either the amenities introduced by plant owners (i.e. often plants have to provide local public goods such as recreational parks to ease site acceptance) or disvalue the plant due to the associated nuclear risks or eyesore of the power plant itself. In order to estimate the (a) net effect of the treatment, we follow the method by Davis (2011), controlling mainly for pre-treatment differences. In order to estimate the

[3]A collection of the decennial censuses from 1970 to 2000 with all values normalized to 2000 tract geographies.

[4]This amenity value can also be considered the willingness to accept (WTA) the introduction of a plant aside from the labor market improvements that are also valued by local residents. This WTA is not nationally representative, however. The populations in areas with the potential for placement are different from the rest of the US.

amenity value, we follow the differences-in-differences (DID) framework as it allows us to isolate the amenity value from the evolution of home values shared both near and far from plants, the basic differences between these two subgroups, as well as the effects from sorting.

In our setting, the estimation of credible treatment effects is challenged because of the non-random geographical placement of nuclear power plants. We find that the geographical areas (tracts) surrounding nuclear power plants have very unique characteristics that are difficult to match to a sufficient number of control tracts. We address this problem with two methods: first we apply a number of subsampling techniques to select the control set of census tracts to be comparable to the set of treated tracts according to pre-treatment observables. Secondly, we apply propensity score weights. This technique weights control observations according to their likelihood of hosting a nuclear power plant. Our identifying assumption is that conditional on subsampling and weighting, the control observations would have developed similarly to the treatment observations in the absence of the intervention. Our results are divided across four separate treatments: (1) plant construction; (2) successful plant placementfrom the subset of plants constructed in the 1970s that eventually produced commercial energy; (3) plant closures; and (4) the Chernobyl accident.

With respect to the first treatment, we find that the construction of nuclear plants induces net outward migration as measured by population density (which decreases by 930 people per square mile in response to treatment) and vacancy rate (which increases by 1.5 percentage points). The net outward migration, however, is not uniform across characteristics of the labor force. We find that the overall skill and education level of the labor force increases, presumably from new residents such as engineers and technicians building the reactor. Overall, nuclear construction, partly by inducing this net outward migration, leads to reductions in housing values within the plume exposure pathway of the nuclear power plant.

Surrounding the sample of successfully placed plants, there is no evidence that plant construction leads to a change in the overall population level: changes in population density and vacancy rates are insignificant. The inward migration of educated, highly skilled workers

from out of town is sufficient to bolster the overall population level surrounding successfully placed plants. In turn, labor market effects are also stronger: the effect on housing values and household incomes is insignificantly negative and unemployment rates significantly decrease. This is in contrast to the more adverse effects discussed above around plant construction when the analysis includes areas surrounding all plant construction and is not limited to only those plants that eventually produce commercial energy. In the case of construction surrounding the construction of successfully placed plants, there may have been unobservables that signaled the plant to be a more credible employer. Greater credibility could lead to these stronger labor market improvements. However, in response to accident risk around this subset of successful plants, ownership rates significantly decrease within 5 miles, an effect unseen around nuclear plant construction in general. Lower ownership rates demonstrate a reduced willingness to make long-term investments in areas subject to potential accidents, however rare, into the future.

In terms of our nonmarket amenity estimation, our difference-in-difference regressions on housing values, controlling for the labor market improvements, show that homeowners around plants experience a negative externality imposed by the plants placement. This is likely due to unsightliness or nuclear accident risk associated with living near the plant. This disamenity value is even stronger in areas within 5 miles of the plant than within the greater 10-mile plume exposure pathway.

Our estimates of the treatment effect of plant closures in the 1990s reinforce the conclusions drawn from plant placements. We find that generally plant closures reverse the original effects of plant placement. In response to closures, the net effects on housing values, ownership rates, and population densities are positive. Reinforcing the consistent reversion of placement effects, our DID amenity value regressions of plant closures reverse the effect of the DID regressions evaluating successful plant placements.

Finally, the treatment of the 1986 Chernobyl accident is analyzed on the census tracts surrounding plants that consistently produced commercial nuclear energy throughout the

1980s (hence comparing 1980 to 1990 values). We find that the accident in the Ukraine further lowered the percentage of owner-occupied homes and reduced housing values around US nuclear power plants. The Chernobyl accident, however, shows no significant impact on the amenity value of nuclear power plants.

The rest of this paper is divided into five parts. The following section discusses the previous literature on the effects of nuclear power plants on local markets. Section 3 discusses the data used in the current paper. Section 4 details the different sets of nuclear power plants available to investigate the impact of major nuclear power plant events and treatment definitions. Section 5 reviews the sample of non-treated locations and the procedures used to generate a comparable control to identify the effect of nuclear power plant events. Section 6 describes the regression analysis methods employed to evaluate the net effect of major events in the lives of US nuclear power plants and discusses the analysis results. Section 7 describes the difference-in-difference methodology used to yield amenity values associated with nuclear power plants and evaluates the estimated disamenity values.

1.2 Previous Literature

The previous literature estimating the effect of nuclear power plants has yielded contrary and weak estimates (Clark and Allison, 1999; Clark et al., 1997; Folland and Hough, 2000; Gamble and Downing, 1982; Nelson, 1981). In addition, non-systematic selections of control areas in the face of non-random treatment assignment make it hard to argue the estimated treatment effect is not due to pre-treatment differences in growth patterns. Finally, the sorting of individuals by preference along the distance gradient from power plants obscures the welfare impact. Galster (1986) discussed the time path one might expect from introducing a plant to an area. He described the endogenous sorting of individuals along a distance gradient to the plant according to their preferences. Assuming some intraurban mobility, non-indifferent households increase their bids on locations safely distant from externalities and lower their bids for proximate locations until the price differential exactly compensates

for the value placed on those externalities. Over time and assuming interurban mobility, indifferent households seeking a bargain or access to new jobs would begin to bid up the nuclear-proximate area and the price differential would disappear if the number of indifferent households exceeds the supply of housing within this zone. In short, the individuals nearest the plants may be risk neutral (as evidenced by their willingness to accept the nuclear power plant in the first place) and so are unaffected by the nuclear power plant. This may not be the case for the average individual.

Clark et al. (1997) analyze property sales data between 1990 and 1994 for the 25-mile area surrounding two California nuclear power plants, Rancho Seco near Sacramento and Diablo Canyon near San Luis Obispo. Using year, suburb, and city fixed effects together with an impressive list of spatial and demographic covariates, the authors find evidence of housing price premiums within 23 miles of Diablo Canyon and 11 miles of the Rancho Seco plants. After these boundaries the housing values increase with distance. The authors seem dissatisfied with the results, when digested with surveys reporting high percentages of the population who would be unwilling to live next to nuclear power stations. However, omitted variable bias and the relocation of households according to individual preferences or risk assessments could explain their results, while Metz (1994) offers numerous arguments invalidating the survey results on this topic (see Appendix Section B).

Clark et al. (1997) do not capture what may make proximity to these particular nuclear power plants appealing. The Diablo Canyon plant is tucked away in the hills of central coastal California close to state parks and estuaries making it a popular place for luxurious private homes. Similarly, the municipality that owns the Rancho Seco plant developed park and picnic spaces surrounding the property. Combined with intuition on endogenous sorting that would lead us to guess that the nearby residents are composed of individuals with limited risk aversion to potential nuclear accidents, the u-shaped price gradient is unsurprising.

In a follow-up to their 1997 paper, Clark and Allison (1999) evaluate the adverse perceptions of spent nuclear fuel siting that may be capitalized into home prices by using the five

years of data surrounding a 1991 announcement to relocate spent nuclear fuel at the Rancho Seco plant to dry storage, a cost saving move after the plants shutdown in 1989. The authors test the significance of a continuous time variable, distance to the closed nuclear plant, dummies for visual cues, and the prevalence of articles written about the nuclear power plant after the announcement to move spent fuel but before the sale of a home, and a full set of interactions and control covariates. The reported results are weak, finding that housing prices increase on average by 4.3% per mile distanced from the plant, with a decreasing effect over time. The plant visibility dummy is perversely found to be positively significant, likely due to omitted variables. The announcement to move spent nuclear fuel, even when interacted with distance and visual cues as well as the variables capturing reported news about the plants, is insignificant.

While the above analyses of home price yielded positive distance gradients (within a given distance), Folland and Hough (2000) find the opposite in their analysis of agricultural land values. Using a panel dataset of eight time periods spanning 1945 and 1992 where the unit of observation is a Basic Trading Area (BTA), an area dominated by similar groups of products like grains or livestock, they estimate the significance of a nuclear presence dummy variable and time since initial operation to agricultural land price. The authors control for the value of the product, population density (to account for encroaching housing markets on agricultural land), land size, distance to major trading centers or ports, and BTA fixed effects. To test the idea that nuclear power plants are placed in less-populated, cheap-land areas, the authors create three so-called ghost reactor variables. The first is much like a reverse DID, an indicator variable equaling one in years prior to the installation of a proximate nuclear power plant and zero after installation or for BTAs without eventual power plants. The second ghost reactor variable captures the announcement effect by equaling one in the year the contract is published. The third ghost reactor variable equals one for every year prior to the contract date. The authors claim that if the first and third variables are negative while using the second to control for the announcement effect, then the cheap-land

hypothesis is supported and the negative externality of power plants is merely the product of the screening by placement decisions. Folland and Hough (2000) find that a significant negative effect remains after controlling for the cheap-land hypothesis with the inclusion of the ghost reactor variables. In addition, the inclusion of these ghost variables reverses the sign on the time-since-installation variables. Without the ghost variables, this effect was positive, possibly supporting the idea that the installation of a nuclear power plant boosts growth in these areas through a larger tax base or labor market improvements. Including these ghost variables (installation effect and cheap-land placement decisions) reversed the nuclear time coefficient suggesting that land prices continue downward after installation.

Gawande and Jenkins-Smith (2001) take a different approach to revealing the negative externalities associated with nuclear power plants by performing a hedonic analysis using well-publicized spent nuclear fuel shipment routes as a treatment. The sample of individual housing sales transactions covers Aiken, Berkeley, and Charleston Counties of South Carolina for four shipments from 1994 to 1996. To test for the effect of the nuclear shipments on housing values disaggregated by county, the authors regressed housing values on distance-to-route variables and many pre- and post-shipment dummies. The disaggregation by counties is especially interesting because of survey results, which showed that Berkeley residents placed the risk of being affected by an accident significantly lower on a 0 (not likely) to 10 (certain) scale. Authors speculate this is due to the experience of Berkeley county residents with their own nuclear power plant, an idea critical to the analysis presented in our paper. For the more urban Charleston County, results show a positive coefficient on interactions between distance to the rail shipment route and pre-treatment dummies. This suggests that proximity to rail alone is a negative amenity, even before shipment. Post-shipment dummies increased, showing the negative impact of proximity to the railways along the shipment path, an effect that persists over time. For the rural nuclear plant containing Berkeley County, distance and post-treatment time dummies were insignificant. This either supports the idea that Berkeley residents are less risk averse to nuclear materials or that in this area nuclear

waste is much less stigmatized.

Aside from Folland and Hough (2000), none of the published literature focuses explicitly on the effect of plant placements. Even then Folland and Hough (2000) focus on effects to agricultural land values, not residential. Some published works do focus on the final topic of our paper, the effect of commercial nuclear accidents. These works however focus on the Three Mile Island accident and that plants surrounding area.

The earliest, Nelson (1981), regresses the housing values in two towns near the Three Mile Island (TMI) Nuclear power plant on an indicator for post-accident sale, a quarterly time variable, and the interaction of the two to capture changes in the rate of housing price growth or decline after the accident. Finding no significant change in housing prices due to the TMI accident, Nelson extends the analysis, providing t-tests for the difference between the average housing price of a 5-mile area around TMI after the accident to a counterfactual average housing price calculated using the appreciation rates from two control areasthe 20-mile area surrounding TMI, and Williamsport, a city 100 miles from TMI. Again, Nelson finds no significance. His regressions suffer from small sample sizes and the lack of a justifiably appropriate control group. In the t-tests, he neither controls for differences in treatment and control areas nor does he satisfactorily argue that those areas can be used as control groups given the endogenous nature of nuclear power plant placement.

The analysis of Gamble and Downing (1982) on the Three Mile Island Accident was more rigorous. Using a sample of 583 single family homes within 25 miles of the plant and 112 homes in Williamsport, Gamble and Downing perform a difference-in-difference analysis with indicator variables for near versus far from plant, before versus after accident, and the interaction of the two. Again, no significance was found associated with the accident even when the samples were broken down by north, south, east and west quadrants around the plant to test for any downwind effects. An argument for the lack of significant negative effects from the accident is the counteracting influx of clean-up crews and technicians that might have increased the demand for housing in the area. However, the breakdown by quadrant

also reflected variation in commute times to the plant. Living in some quadrants requires the inconvenient crossing of the Susquehanna River. Even so, no differences in the effect of the accident were found across quadrants. Gamble and Downings analysis still suffers from endogeneity bias as those who located near the TMI plant before the accident may have lower risk aversion and interpreted the aftermath of the accident as a disaster well-handled and kept in control. It has also been speculated that local populations would have found the area to be safer after the accident because the reactors were immediately shutdown and remained off during the period covered by the Gamble and Downing study.

As discussed, the weak estimates and contrary results of the above literature can be explained by the confluence of identification issues. Due to endogenous sorting, cross-sectional estimates using distance gradients from pre-existing plants are likely to reflect the preferences of individuals who are less likely to view nuclear power plants as sources of risks and instead value the access to high-tech jobs, the plants contributions to the local tax base, or other correlated unobserved features. Cross sectional studies also ignore the potential bias introduced by the non-random site selection process. The previous literature tends to focus on one or two nuclear power plants, leaving their results not easily generalized to the population as a whole even while appropriately controlling for omitted variables. When time dimensions are considered, the authors make little empirical justification as to the appropriateness of their control groups.

1.3 Data

Our dataset uses average reported housing values, housing structure, housing market, demographic, and employment variables from the Neighborhood Change Database (NCDB), a collection of the decennial long form censuses from 1970 to 2000. Each years values are normalized to the 2000 census tract geographies. The NCDB is the most comprehensive national dataset available with the variables needed for our analysis. No other national dataset provides housing values in 1970, the last decade during which new nuclear plant

construction took place in the US.

To complement the census data, we created spatial explanatory variables important to predicting both housing values and nuclear power plant placement. These variables include distance to the nearest city center with a population greater than 25,000[5] and distance to sufficiently large water sources for safe and consistent nuclear power plant operation. Water distance also serves as a proxy for proximity to recreation areas and environmental amenities. See Appendix D, Section 2.1 for a discussion of nuclear power plant siting policy during the 1970s and the estimation of these spatial variables.

We compiled important power plant dates and geographic coordinate information from the NRC. Together with census tracts shapefiles from the National Historical Geographic Information System (NHGIS) we identify treated tracts and create the treatment group and control sample.

1.4 Treatment Group

To identify the US nuclear power plants relevant to estimating the effect of our three major treatmentsplant placement, the Chernobyl accident, and plant closuresour study distinguishes four sets of nuclear power plants available and relevant to testing these three events in the lives of US plants.

1.4.1 Plant Placement

We use two sets of nuclear power plants to study the effect of plant placement. The first set is composed of all plants placed after 1971. The control period is then 1970 and the first observations after the event are drawn in 1980, our treatment period. We define the date of placement as the date of earliest construction. While an ideal treatment group to estimate the effects of plant placement would be composed of tracts near nuclear power

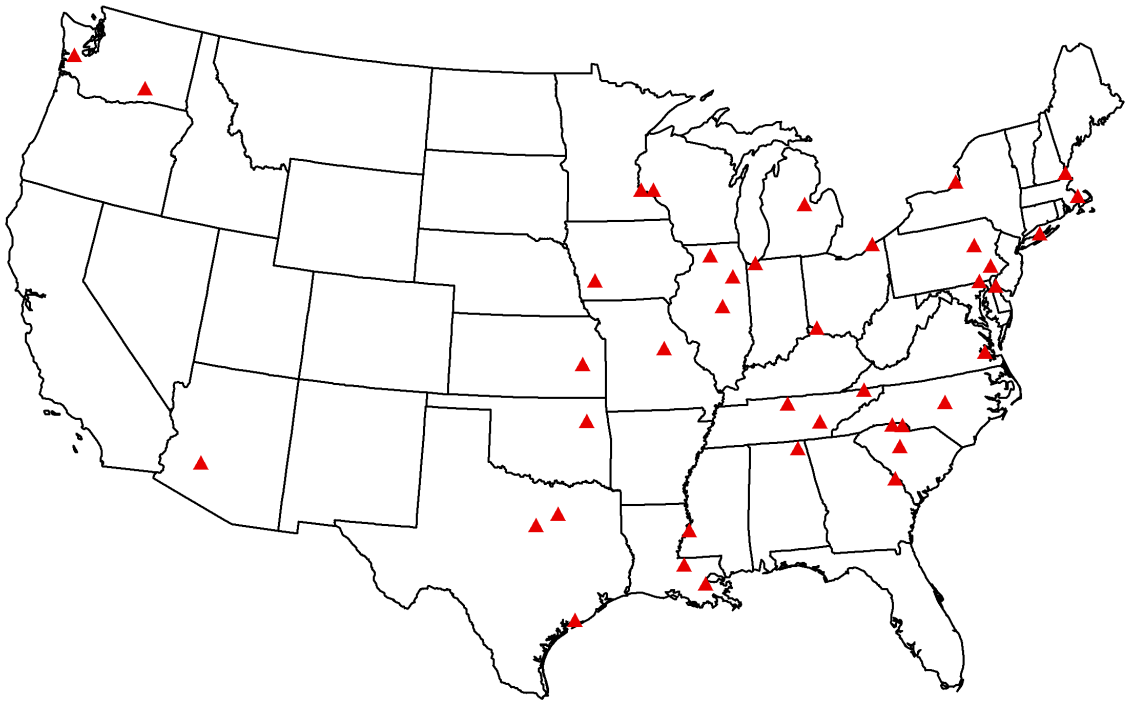
[5]An applicant for a commercial nuclear power construction permit needed to define the population center distance of the proposed site, defined as the distance from a reactor to the nearest boundary of a populated center amounting to 25,000 residents.

plants that were originally proposed, constructed, and began operating between the control and treatment years, no nuclear power plants in the US match this description due to the long and uncertain licensure process and the construction itself. Rumors about a proposed site usually circulated before construction and construction itself more often than not took over a decade. However, Kiel and McClain (1995) break the time before and after the introduction of a landfill incinerator near Boston into various stages: pre-rumor, rumor, construction, online, and ongoing operation. Their results show little to no effect in the rumor stage, the largest effect in the construction and initial operation stage, and smaller persistent effects after ongoing operation out to seven years. Households might not have treated nearby site proposals as credible threats because energy companies often proposed multiple potential locations for the same reactor plan. Also, the newness of commercial nuclear power regulation contributed much uncertainty to the success of any given proposal. Therefore, we believe earliest construction and the financial commitment it represents constitutes a credible threat that a nuclear power plant will be introduced to an area.

There are 40 plants in the US, which began their construction after December 31, 1971 (almost two years after the April 1st enumeration date of the 1970 census) without terminating construction or stopping operation, at least until after 1980. We will hereafter refer to this set of nuclear power plants as the plant placement set. Figure 1 provides a map of these 40 plants in the plant placement set.

The second set of plants we use to study the effect of plant placement is a subset of the plant placement set. Of this set of 40, only 28 plants went on to operate and produce commercial nuclear energy for nearby utilities. From a data collection perspective, there is nothing observable from the 1970s that distinguish these successfully placed plants from the broader plant placement set. However, there may have been something about the speed of construction or the attitudes of plant owners and locals observable at the time that made placement around these plants more credible. We expect a priori that areas around these plants were more strongly effected than around unsuccessful placements. We will hereafter

Figure 1.1: Nuclear Plants Placed After 1971 and Before 1980



refer to the set of nuclear power plants that were placed after 1971 and eventually produced commercially as the successful plant placement set. Figure 2 provides a map of the plants in the successful plant placement set.

1.4.2 Plant Closures

Plant closures provide an opportunity to test the reversion of the placement treatment. Most of the plants to cancel their operations did so during the 1980s after construction or operation costs together with local backlash from the Three Mile Island and Chernobyl accidents became untenable. Unfortunately, the timing of these accidents concurrent with closures makes it near impossible to separate one effect from the other. Therefore, we estimate the effect of nuclear power plant closures with the set of plants that closed after 1990 but before 2000. Hereafter we will refer to this set of plants as the plant closure set. Figure 3 illustrates the locations of the plants in the plant closure set.

Figure 1.2: Nuclear Plants Placed After 1971 and Before 1980, which Successfully Operated

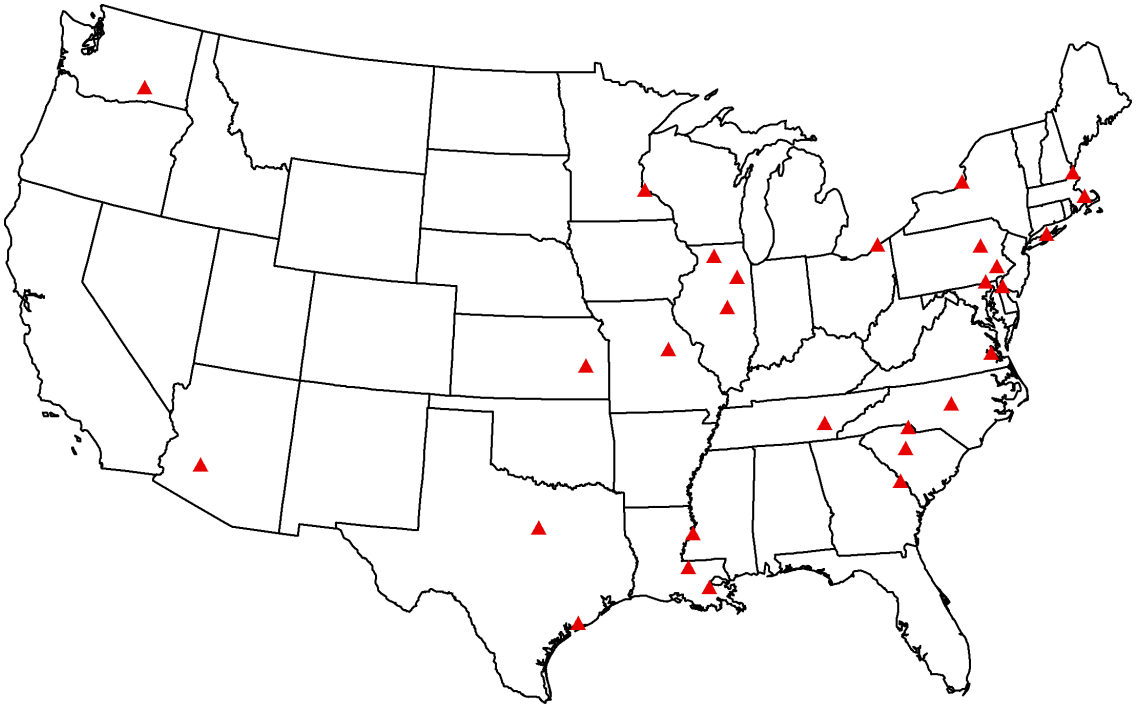


Figure 1.3: Nuclear Plants Closed in the 1990s

1.4.3 The Chernobyl Accident

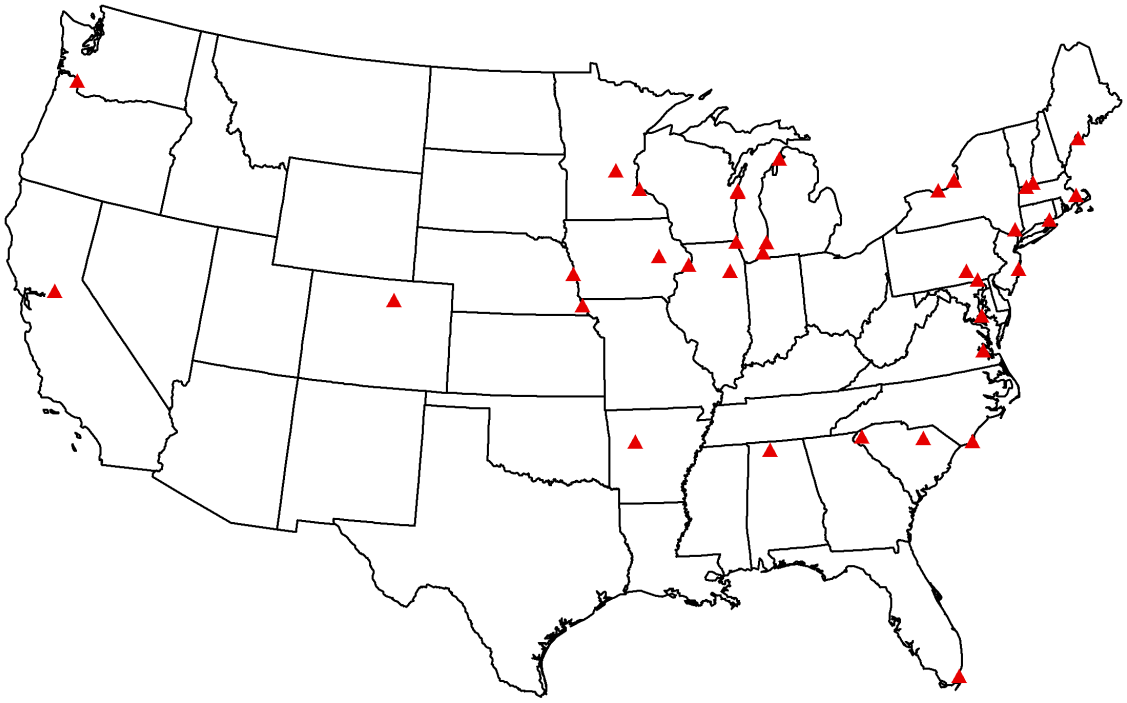
As a large international disaster ten times larger than the Fukushima accident, Chernobyl constitutes an information shock to the US housing market with respect to the potential, horrific, damage of nuclear disasters. The massive explosion and spread of radiation led to 31 deaths due to radiation sickness within the first few weeks. In contrast, the Three Mile Accident produced no known deaths. While the Three Mile Island accident certainly heightened concerns, many interpreted the details to demonstrate the effectiveness of safety procedures. Within moments after the initial malfunction, which drained the cooling water from the containment structure, the Three Mile Island reactor shut itself down by automatically dropping control rods to stop the flow of neutrons that sustained the chained reaction. Even without this safety feature, because the cooling water was also used as a moderator in the reaction, without it the reaction would have slowed down on its own. For these reasons, we view the Chernobyl accident as a stronger information effect over the kinds of damage possible after a large commercial nuclear accident.

Thirty-seven plants began commercial operation before 1980 and continued operation through 1990. This provides an opportunity to test the effect of the Chernobyl disaster in 1986 on the risk perceptions of economic agents surrounding these plants without confounding effects from plant openings or closures. Hereafter we will refer to the set of US nuclear power plants that began operation before 1980 and continued operations through 1990 as the accident set. Figure 4 provides the locations of the plants in the accident set.

1.4.4 Defining Treatment

For each set, census tracts are allocated to the treatment groups if they are sufficiently near the plants within the set. We define what it means to be sufficiently proximate and so in the treatment group ($N_i = 1$) according to two definitions. The first assigns treatment to

Figure 1.4: Nuclear Plants Operating Before 1980 and Through 1990



a tract if the Plume Exposure Pathway[6] (10-mile buffer zone) of any one or combination[7] of nuclear power plants in the set covers greater than 80% of the tracts areas. The second narrows the treatment sample by approximately half by assigning treatment to a tract if 5-mile buffer rings around the set of plants cover more than 50% of the tracts area. See Appendix D, Section 1 for a discussion on the definition of treatment used.

There are two major features of our dataset that reduce the number of nuclear power plants that are available for our investigation from the numbers discussed in the previous subsections. Namely, these features are the clustered nature of nuclear power plants and the reduced coverage of pre-1990 census data.

---

[6]After the Three Mile Island Accident in 1979, the U.S. Nuclear Regulatory Commission (NRC) instituted a ten-mile evacuation zone from plants in the case of an emergency, defining this zone as the Plume Exposure Pathway.

[7]Nuclear accidents are such rare events and yet the potential damages so great as to be classified as uninsurable with even unlimited coverage (Hofert and Wüthrich, 2011). It is then reasonable to conclude that rational agents would react little to the greater probably of suffering from a nuclear accident due to being proximate to multiple plants.

The first of these features is the geographical clustering of nuclear power plants. Plants used to estimate the effect of one type of treatment are often in close proximity to members of plants used to test a different treatment. This can interfere with a clean sample. For example, some plants placed in the 1970s are nearby plants placed much earlier and commercially operating by the 1980s. The presence of these plants contaminates the sample with census tracts with nuclear power plant experience during the control year for plant placement. We would expect households to have already responded in part to the presence of nuclear power, perhaps by moving away to avoid risk or moving closer to access new engineering jobs. To avoid contaminating the treatment and control groups with this early sorting or the confounding presence of other events, we drop tracts within 20 miles of nuclear power plants not in the relevant set for the treatment intended. The blue rings in Figures 5 through 8 demonstrate these 20-mile sample exclusion areas.

Figure 1.5: Nuclear Plants Placed After 1971 and Before 1980

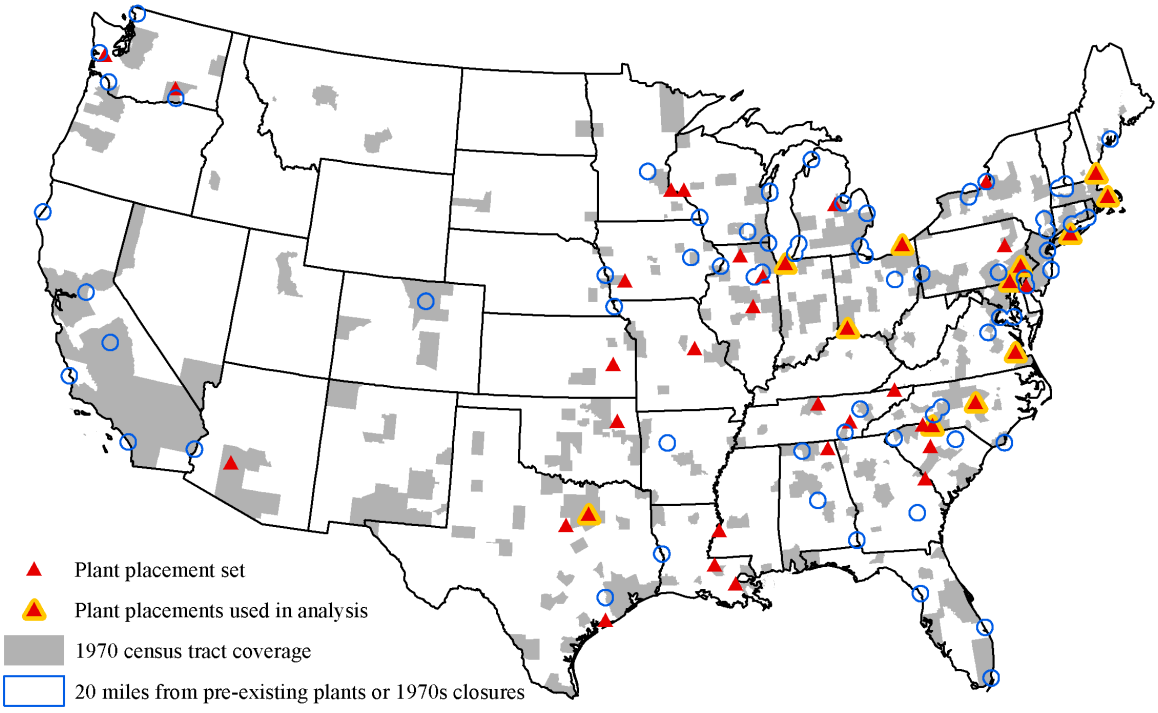


Figure 1.6: Nuclear Plants Placed After 1971 and Before 1980, which Successfully Operated

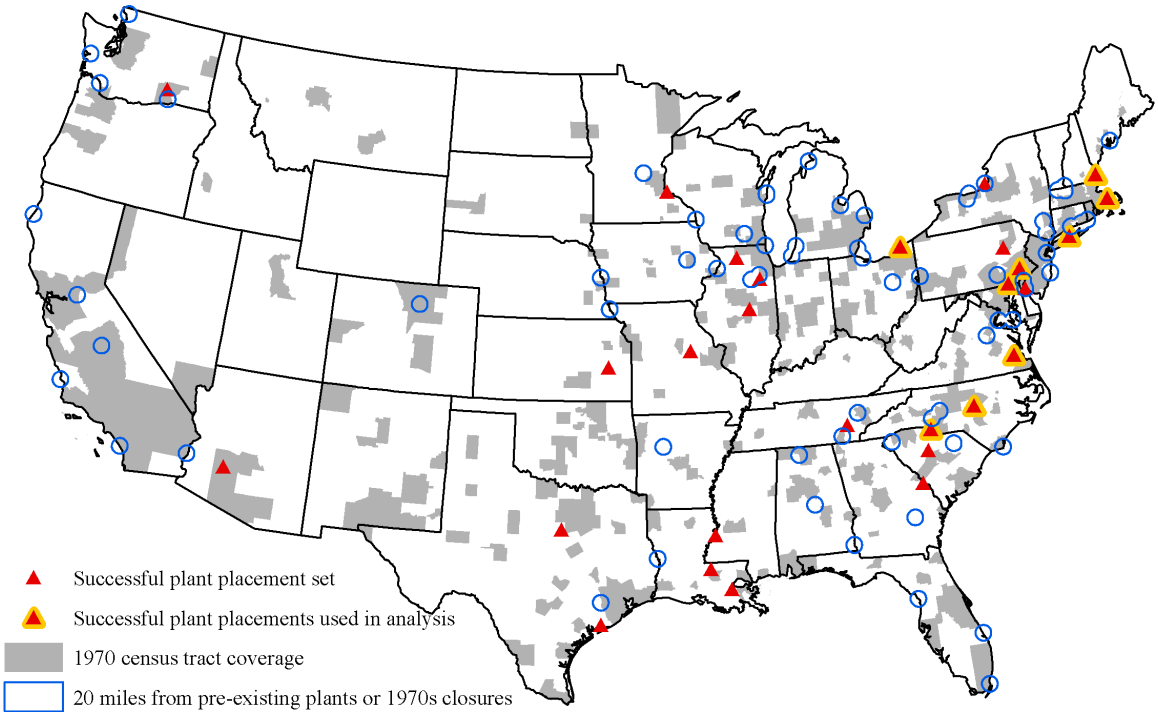


Figure 1.7: Nuclear Plants Closed in the 1990s

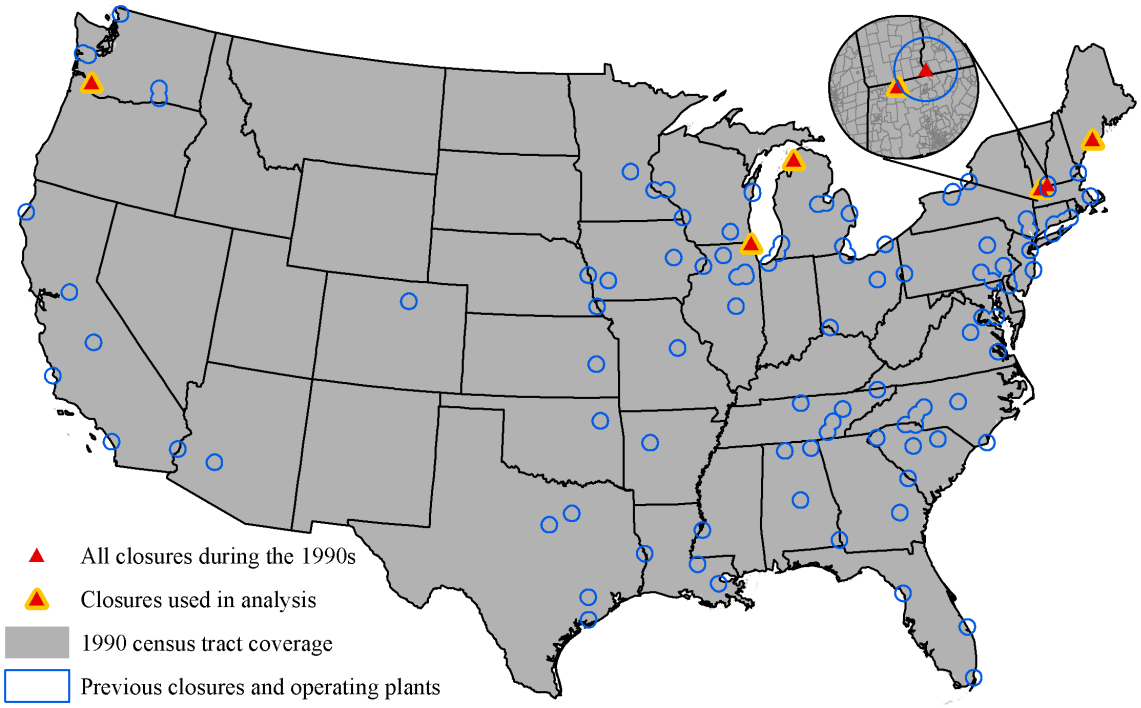
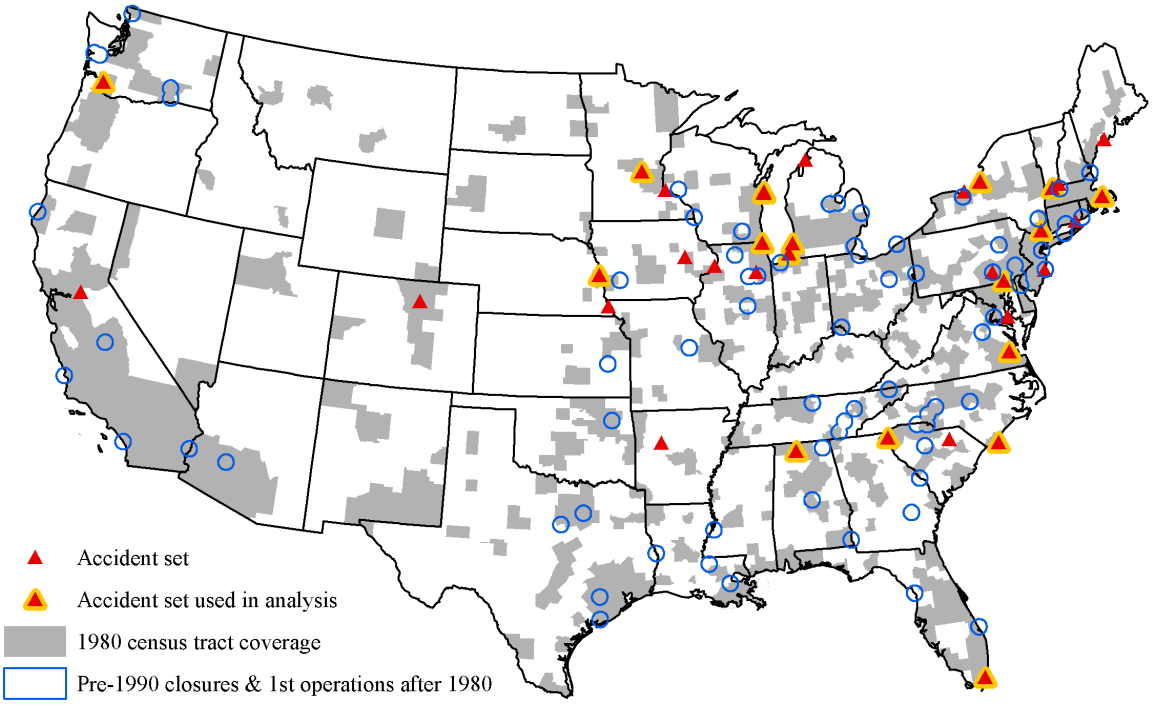


Figure 1.8: Nuclear Plants Operating Before 1980 and Through 1990



The second feature of our dataset that reduces the number of nuclear power plants available for our investigation is the limited geographical coverage of census tracts in the 1980s and 1990s. In many of the areas surrounding nuclear power plants, population densities did not merit defining census tracts at all, and so local demographic and labor market data is unavailable. The geographical coverage of the decennial data used as the control year for each of the treatments studied is illustrated in Figures 5 through 8.

Even when defined, census tracts in less populated areas can be large enough that even ten-mile buffer rings surrounding nuclear power plants cannot sufficiently cover the tract area such that enough of the population can be considered treated. This, together with applying the necessary sample exclusions to deal with the confounding presence of treatments from other nuclear power plant sets, reduces the number of treated tracts and the number of plants remaining in each of the five sets. Table 1 provides the resulting treatment group sample sizes by plant for each of the nuclear power plant sets.

Table 1.1: Treatment group sample sizes by nuclear power plant set

		Earliest Construction	Earliest Commercial Operation	Closure Date	10-mile	5-mile
Placements in the 1970s						
Yellow Creek	MS	1978		1984	195	68
Zimmer	OH	1972		1984	114	13
Shoreham	NY	1973	1989	1989	28	8
Perry	OH	1974	1991		20	8
Catawba	SC	1974	1989		18	3
Surry	VA	1972	1976		15	0
Limerick	PA	1974	1990		14	4
Bailly	IN	1974		1981	11	3
Peach Bottom	PA	1972	1978		3	0
Shearon Harris	NC	1978	1991		2	0
Pilgrim	MA	1972	1976		1	0
Seabrook	NH	1976	1994		1	0
1990s Closures						
Zion	IL	1968	1973	1997	64	21
Trojan	OR	1970	1975	1992	14	0
Maine Yankee	ME	1968	1973	1996	7	0
Big Rock Point	MI	1953	1962	1997	3	3
Yankee Rowe	MA	1954	1963	1991	1	0
Plants to test Chernobyl Accident						
Zion	IL	1968	1973	1997	60	21
Indian Point	NY	1962	1978		49	16
Surry	VA	1972	1976		15	0
Duane Arnold	IA	1969	1979		14	1
Nine Mile Point	NY	1965	1973		14	2
Turkey Point	FL	1967	1976		13	0
Trojan	OR	1970	1975	1992	13	0
Donald C. Cook	MI	1969	1979		12	3
Pilgrim	MA	1972	1976		9	3
Monticello	MN	1969	1975		6	1
Peach Bottom	PA	1972	1978		5	1
Palisades	MI	1966	1976		4	3
Oconee	SC	1967	1977		4	1
Fort Calhoun	NE	1966	1977		4	1
Browns Ferry	AL	1967	1978		4	0
Point Beach	WI	1971	1974		3	0
Brunswick	NC	1969	1979		2	0
Kewaunee	WI	1967	1978	2013	2	2
Yankee Rowe	MA	1954	1963	1991	1	0

Notes: Plant placements with a non-missing commercial operation date are part of the successful placement set.

1.5 Control Group

Our first step in cleaning the set of non-treated census tracts to be used as control observations is to remove any tracts that could be considered partially treated. To avoid a dichotomous distinction where tracts situated 9.9 miles away are considered treated yet census tracts 10.1 miles away are not, we exclude tracts within a 20-mile buffer from nuclear power plants in the treatment set that do not met the treatment criteria, effectively dropping donut shaped rings of tracts around treated areas. The remaining non-treated tracts will hereafter be referred to as baseline control tracts. Figure 9 illustrates this specification around select nuclear power plants in the successful plant placement set.

The next step is to ensure that the set of tracts used as counterfactuals for the treatment groups are appropriately comparable. The areas surrounding nuclear power plants are very different from the rest of the US on average. This is most evident for population density. Figure 10 illustrates the distributions of population density in 1970 both within the whole sample of tracts and those belonging to the treatment group of the plant placement set. Very rarely were plants placed in areas with population densities above 30 people per square mile. In the whole sample, this population density is more common. In contrast, the distributions of vacancy rate in the treatment and control groups are similar even at the tails.

When comparing sample means, tracts within 10-miles of nuclear plants in the plant placement set are 60% as dense as the baseline control tracts on average. Tracts in the treatment sample for the successful plant placement set are less than 35% as dense. They are also closer to water sources and more distant from metropolitan areas, and occupied by a less educated, poorer and less diverse population (see Table 2; Columns 1 and 2).

Figure 1.9: Defining baseline control tracts: Successful placement set

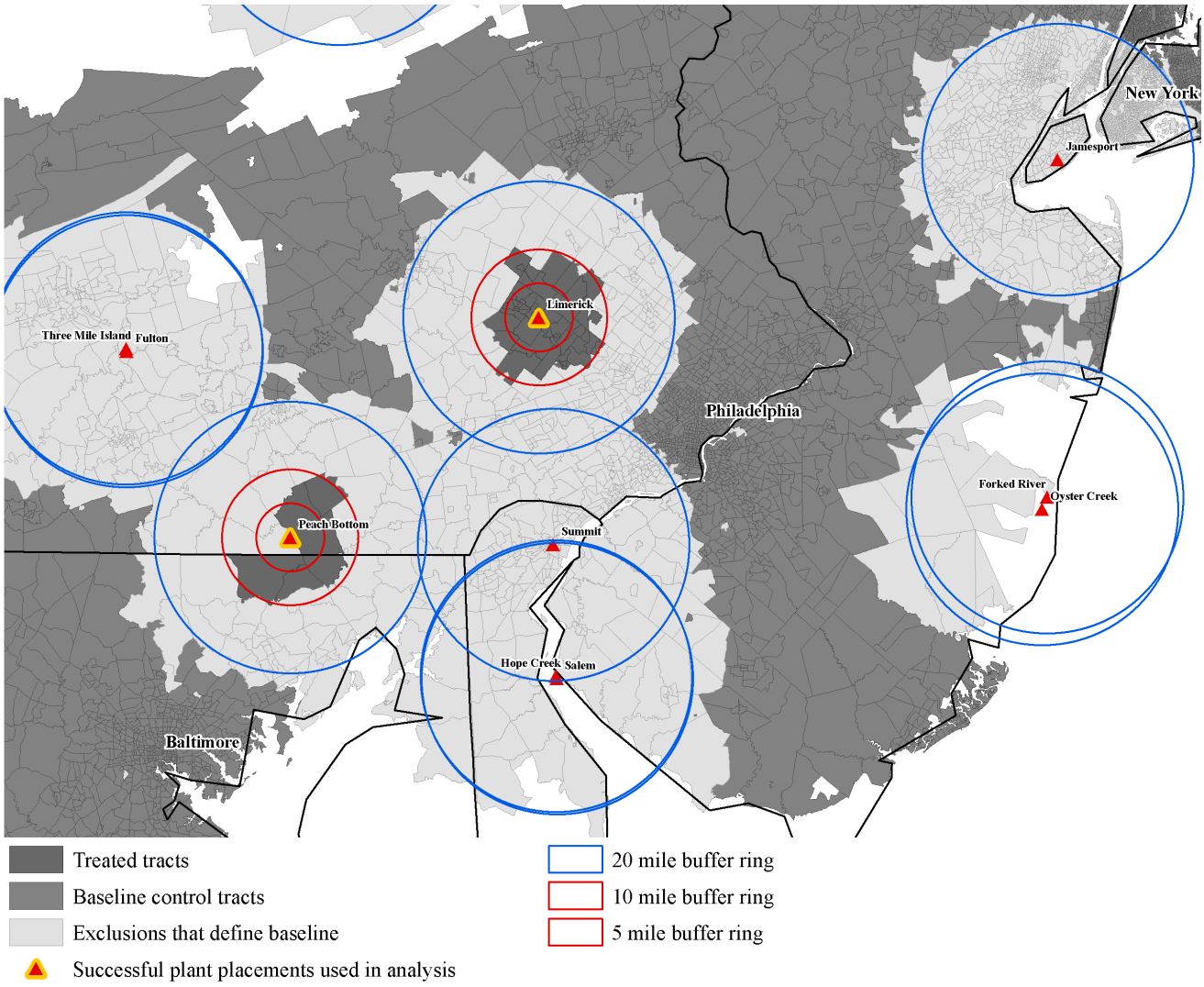
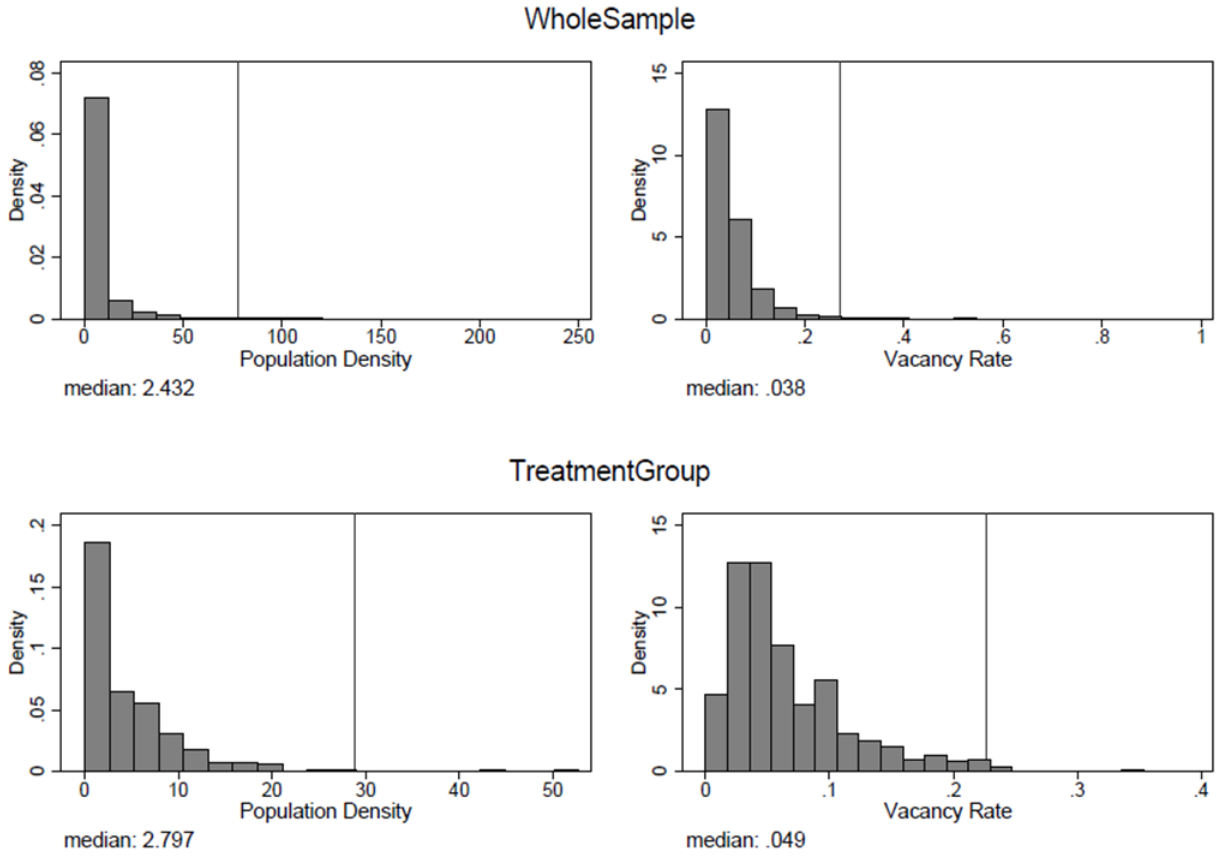


Figure 1.10: Distributions of Population Density and Vancancy Rate for Whole Sample and Treatment Group of Plant Placement Set



Notes: Vertical lines correspond to sample medians.

Table 1.2: Sample means of the treatment and baseline control groups for each of the nuclear power plant sets used to evaluate the effect of placement, closures, or the Chernobyl accident.

	Plant placement		Successful plant placement		Plant closures		Chernobyl accident	
	treatment group	control group	treatment group	control group	treatment group	control group	treatment group	control group
Miles to populous city centers	7.7	15.93***	12.79	15.93***	9.04	13.64	13.72	14.01***
Miles to sufficient water source	4.73	7.67***	2.62	7.67***	1.93	8.62***	2.07	7.84***
Pop density (1,000 per mi^2)	4.76	5.9	1.45	5.9***	3.42	4.33*	2.52	4.9***
Out-of-county immigrants (%)	24.2	27.17***	25.29	27.17	20.76	21.15	22.95	26.45***
Black (%)	16.45	9.63***	7.19	9.63	10.76	12.78	8.93	11.64***
Hispanic (%)	4.11	6.3***	1.3	6.3***	8.48	8.47	5.11	7.24**
At least 12 years education (%)	27.94	32.07***	32.88	32.07	32.77	29.97***	37.92	34.21***
Al least bachelors degree (%)	10.07	12.12***	8.97	12.12***	15.54	19.29***	16.42	17.33**
HH income ($1,000s)	10.07	11.03***	10.4	11.03***	36.8	37.33	21.86	21.09
Unemployment rate (%)	3.7	4.43***	3.19	4.43***	6.57	7.12	6.76	6.41
Professional and technical (%)	13.53	15.79***	14.55	15.79	15.36	16.8*	15.79	15.43
Ave housing value ($1,000s)	18.41	21.52***	20.41	21.52**	85.08	103.75***	55.49	59.51***
Vacancy rate (%)	6.5	5.5***	5.42	5.5**	7.63	9.7*	6.16	6.86*
Tenure rate (%)	61.19	67.14***	74.15	67.14***	63.71	65.1	68.27	65.88**
Multi-units structures (%)	31.83	26***	14.87	26***	32.09	26.27*	29.06	29.63
Structures < 5 years old (%)	15.24	18.22**	17.88	18.22	9.29	10.81	14.16	15.45*
Structures < 10 years old (%)	28.1	33***	30.27	33	15.03	20.26**	28.76	28.96

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Migrants are those individuals who have moved to the ares from a different county within the last 5 years.

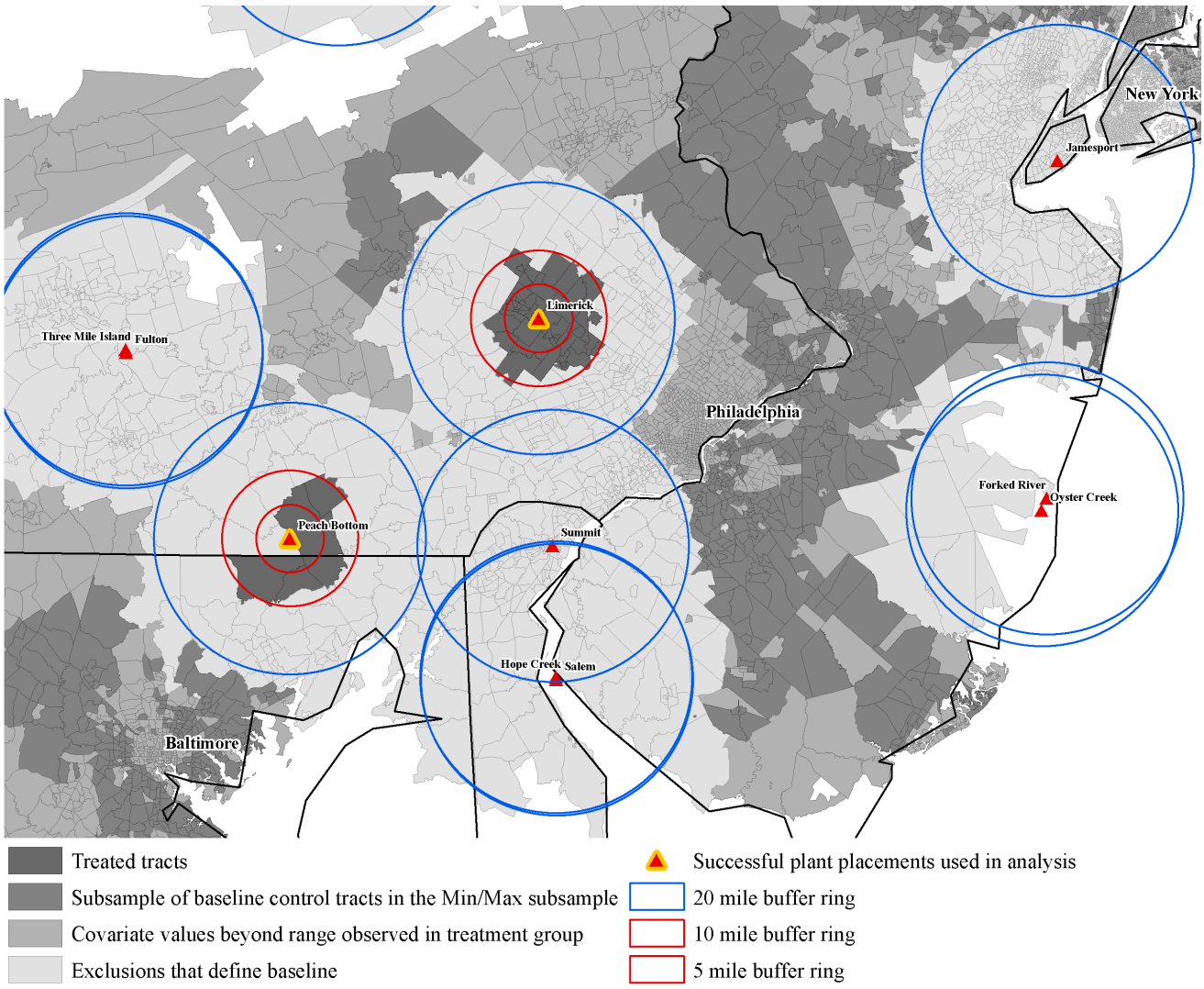
The difference in sample means of individual covariates discussed above provides evidence of imbalance. Balance refers to the degree in which the covariates multi-dimensional distribution of the treatment group resembles the multidimensional distributions of the control group. In other words, it describes the degree to which non-treated observations can serve as counterfactuals to the treated observation. With more than three covariates, it is very difficult to inspect multi-dimension distributions for balance. However, the significant difference in sample means for the majority of covariates serves as evidence that imbalance in this dataset is a problem.

We address imbalance by first subsampling the full set of non-treated census tracts to include only those that are potential candidates for nuclear power plants. We then weight the control observations by the relative odds of treatment conditional on covariate values. In this way, the covariate observations that are more similar to treatment observations in covariate space are weighted more heavily. The subsampling procedure performed also alleviates some of the variability that will be introduced to the final regressions by the estimation of weights (Crump et al., 2006).

We use four different rules, ordered by increasingly aggressive subsampling, to subsample the control group to increase its comparability to the treatment group. The first of these rules is applied only to population density. We use the Blocked Adaptive Computationally Efficient Outlier Nominators (BACON) algorithm applied to population density to identify outliers within the treatment group. With only one variable, this amounts to modeling population density with a χ^2 distribution to identify outliers. We then drop control observations with population densities beyond those of the treatment group outliers. Hereafter we will refer to the sample generated under this rules as the BACON subsample.

The second rule drops control observations with values of any covariate beyond those observed in the treatment groups. This rule will be referred to as generating the Min/Max subsample. Comparing Figures 9 and 11 demonstrates the differences between baseline control tracts and the Min/Max subsample. Notice that most tracts in downtown Philadelphia

Figure 1.11: The Min/Max Subsample: Successful placement set



and Baltimore, for example, have been removed from the set of control tracts used.

The third and fourth rules apply even more aggressive subsampling of the control observations. For rule three we dropped control observations with values beyond the 2^{nd} and 98^{th} percentiles observed among the treatment groups, which defines the 2^{nd} & 98^{th} subsample. The final rule drops tracts beyond the 4^{th} and 96^{th} percentiles, which will be referred to as the 4^{th} & 96^{th} percentile. Even with aggressive subsampling, the control groups remain significantly different from the treated group. Tables 3 through 6 provide sample means for the treatment group within the plume exposure pathway as well as all five possible control group

subsamples: the baseline control tracts, the BACON subsample, the Min/Max subsample, the 2^{nd} & 98^{th} percentile subsample, and the 4^{th} & 96^{th} percentile subsample.

The significant differences that remain after subsampling motivates the additional procedure of creating propensity score weights to be used in our regression analysis. After reducing the sample of control observations to be more comparable to tracts surrounding the nuclear power plant sets, we estimate the probability of treatment conditional on the set of explanatory variables important to the nuclear power plant siting decision. We predict the probability of treatment using logistic regressions on the nuclear proximity treatment variable, N_i $(i = 1, ..., m)$, which takes on the value 1 with probability n_i and 0 with probability $(1 - n)_i$. The binary observations are assumed distributed according to the Bernoulli $(N_i|_i)$ distribution where $_i = \frac{1}{1+e^{-x_i\beta}}$.

Table 1.3: Summary statistics for plant placement set by 10-mile treatment group and control subsamples.

		Control group subsamples			
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	7.70	15.93***	5.33	5.29	4.87	4.71**
Miles to sufficient water source	4.73	7.67***	4.93	4.88	4.19***	4.05***
Pop density (1,000 per sq-mile)	4.76	5.9	5.05	5.63***	4.84	4.47
Living in a diff county 5 years ago (%)	24.20	27.17***	23.82	23.76	22.82	22.23**
Black (%)	16.45	9.63***	8.7***	9.31***	8.63***	7.43***
Hispanic (%)	4.11	6.3***	5.29	5.4	3.72**	2.86***
At least 12 years education (%)	27.94	32.07***	32.83***	32.66***	33.38***	33.47***
Al least bachelors degree (%)	10.07	12.12***	11.51*	11.41*	10.24	9.69*
HH income ($1,000s)	10.07	11.03***	11.01***	10.95***	10.49*	10.35
Unemployment rate (%)	3.70	4.43***	4.44***	4.49***	4.38***	4.14***
Professional and technical (%)	13.53	15.79***	15.56***	15.44***	14.91***	14.62**
Average housing value ($1,000s)	18.41	21.52***	21.15***	21.1***	19.69***	19.04**
Vacancy rate (%)	6.50	5.5***	4.82***	4.84***	4.46***	4.31***
Tenure rate (%)	61.19	67.14***	66.86***	66.03***	66.04***	67.03***
Multi-units structures (%)	31.83	26***	25.45***	26.53***	26.86***	26.12***
Structures $<$ 5 years old (%)	15.24	18.22***	15.9	15.64	14.47	14.24
Structures $<$ 10 years old (%)	28.10	33***	29.7	29.26	27.27	26.56**

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

Table 1.4: Summary statistics for successful plant placement set by 10-mile treatment group and control subsamples.

		Control group subsamples			
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	12.79	15.93***	5.77***	5.51***	7.11**	7.61
Miles to sufficient water source	2.62	7.67***	3.48***	3.46***	2.95	2.92
Pop density (1,000 per sq-mile)	1.45	5.9***	3.15***	3.73***	1.84***	1.38
Living in a diff county 5 years ago (%)	25.29	27.17*	22.81	22.57	22.14	22.11
Black (%)	7.19	9.63	3.98**	4.12**	3.57**	2.95***
Hispanic (%)	1.30	6.3***	1.62	1.65	0.94**	0.89**
At least 12 years education (%)	32.88	32.07	35.42***	35.24***	35.46***	35.36***
Al least bachelors degree (%)	8.97	12.12***	10.54**	10.63**	8.59	7.72**
HH income ($1,000s)	10.40	11.03***	11.09***	11.11***	10.58	10.45
Unemployment rate (%)	3.19	4.43***	3.89**	3.87**	3.83**	3.59
Professional and technical (%)	14.55	15.79	15.39	15.49	13.86	13.23*
Average housing value ($1,000s)	20.41	21.52***	20.16	20.25	18.84*	18.56**
Vacancy rate (%)	5.42	5.5***	4.09***	4.01***	4.04***	3.98***
Tenure rate (%)	74.15	67.14***	74.11**	73.11***	76.59	78
Multi-units structures (%)	14.87	26***	18.8***	20.63***	15.12***	13.33**
Structures < 5 years old (%)	17.88	18.22	17.03	16.3*	17.96	18.59
Structures < 10 years old (%)	30.27	33	30.96	29.82	31.98	32.9

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

Table 1.5: Summary statistics for plant closure set by 10-mile treatment group and control subsamples.

		Control group subsamples				
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	9.04	13.64*	11.85	11.85	12.62	6.88
Miles to sufficient water source	1.93	8.62***	3.91***	3.91***	3.01***	2.88***
Pop density (1,000 per sq-mile)	3.42	4.33**	2.55***	2.55***	2.43***	2.71**
Living in a diff county 5 years ago (%)	20.76	21.15	19.18**	19.18**	19.27*	19.8
Black (%)	10.76	12.78	10.12	10.12	8.54	7.86*
Hispanic (%)	8.48	8.47	4.03***	4.03***	3.82***	4.02***
At least 12 years education (%)	32.77	29.97***	33.48	33.48	33.74	33.96*
Al least bachelors degree (%)	15.54	19.29***	15.46	15.46	14.78	14.99
HH income ($1,000s)	36.80	37.33	33.54***	33.54***	32.94***	33.93***
Unemployment rate (%)	6.57	7.12	6.29	6.29	6.35	6.03
Professional and technical (%)	15.36	16.8**	15.69	15.69	15.46	15.6
Average housing value ($1,000s)	85.08	103.75***	76.35**	76.35**	75.02***	79.17*
Vacancy rate (%)	7.63	9.7**	8.05	8.05	8.03	7.15
Tenure rate (%)	63.71	65.1	67.15	67.15	67.08	65.74
Multi-units structures (%)	32.09	26.27**	24.52***	24.52***	24.28***	26.96*
Structures $<$ 5 years old (%)	9.29	10.81	9.5	9.5	9.63	9.94
Structures $<$ 10 years old (%)	15.03	20.26***	18.48*	18.48*	18.62**	18.89**

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

Table 1.6: Summary statistics for Chernobyl accident set by 10-mile treatment group and control subsamples.

		Control group subsamples				
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	13.72	14.01***	6.94***	6.86***	7.14***	6.97***
Miles to sufficient water source	2.07	7.84***	4.58***	4.57***	3.11***	2.5***
Pop density (1,000 per sq-mile)	2.52	4.9***	3.18***	3.35***	2.58	2.2**
Living in a diff county 5 years ago (%)	22.95	26.45***	20.72***	20.63***	19.8	19.85
Black (%)	8.93	11.64***	8.13	8.44	5.23***	4.24***
Hispanic (%)	5.11	7.24***	4.16**	4.21**	3.16***	2.36***
At least 12 years education (%)	37.92	34.21***	37.1**	37.02**	39.01**	39.56***
Al least bachelors degree (%)	16.42	17.33***	15.75	15.75	14.72	14.64
HH income ($1,000s)	21.86	21.09	21.05**	21**	21.02**	21.02**
Unemployment rate (%)	6.76	6.41	6.18***	6.21***	6.07***	6.19***
Professional and technical (%)	15.79	15.43	15.12	15.13	14.89	15.03
Average housing value ($1,000s)	55.49	59.51***	56.6	56.63	53.19	52.49**
Vacancy rate (%)	6.16	6.86**	5.81	5.8	5.09**	4.97**
Tenure rate (%)	68.27	65.88***	69.06	68.67	71.09**	71.68**
Multi-units structures (%)	29.06	29.63	25.04**	25.51**	23.09***	23.01***
Structures < 5 years old (%)	14.16	15.45**	13.11	12.92	13.39	14.05
Structures < 10 years old (%)	28.76	28.96	26.81	26.45	27.37	28.6

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

This allows the probability of treatment to vary according to our set of explanatory variables, X_i, provided in Tables 2. Also included in Xi are state fixed effects to capture the differences in regulation stridency at the state level, the number of individuals under 18 years old per household, the number of individuals over 65 per household, the percent black, the percent Hispanic, and the percents of employed individuals working within various occupational categories. Maximum likelihood estimation is used to find estimates of the parameters, β, and calculate the predicted conditional probability of proximity for each census tract, $\Pr(N_i = 1X)$. Control observations are then weighted by the relative odds of treatment to emphasize tracts similar to the targeted treatment group for the identification of the effect of placement, the Chernobyl accident, or plant closures. The weights are defined by the following scheme.

$$w_i = \begin{cases} \frac{Pr(N_i=1|X)}{1-Pr(N_i=1|X)} & \text{if } N_i = 0 \\ 1 & \text{if } N_i = 1 \end{cases} \tag{1.1}$$

Once weighted, the control group is insignificantly different from the treatment group during the control period as seen by the p-value from t-tests on the difference between the treatment and control group subsamples when weighted by the relative odds of treatment provided in Tables 7 through 10. Figure 12 illustrates these weights applied to the Northeast.

Figure 1.12: The Min/Max Weighted Subsample: Successful placement set

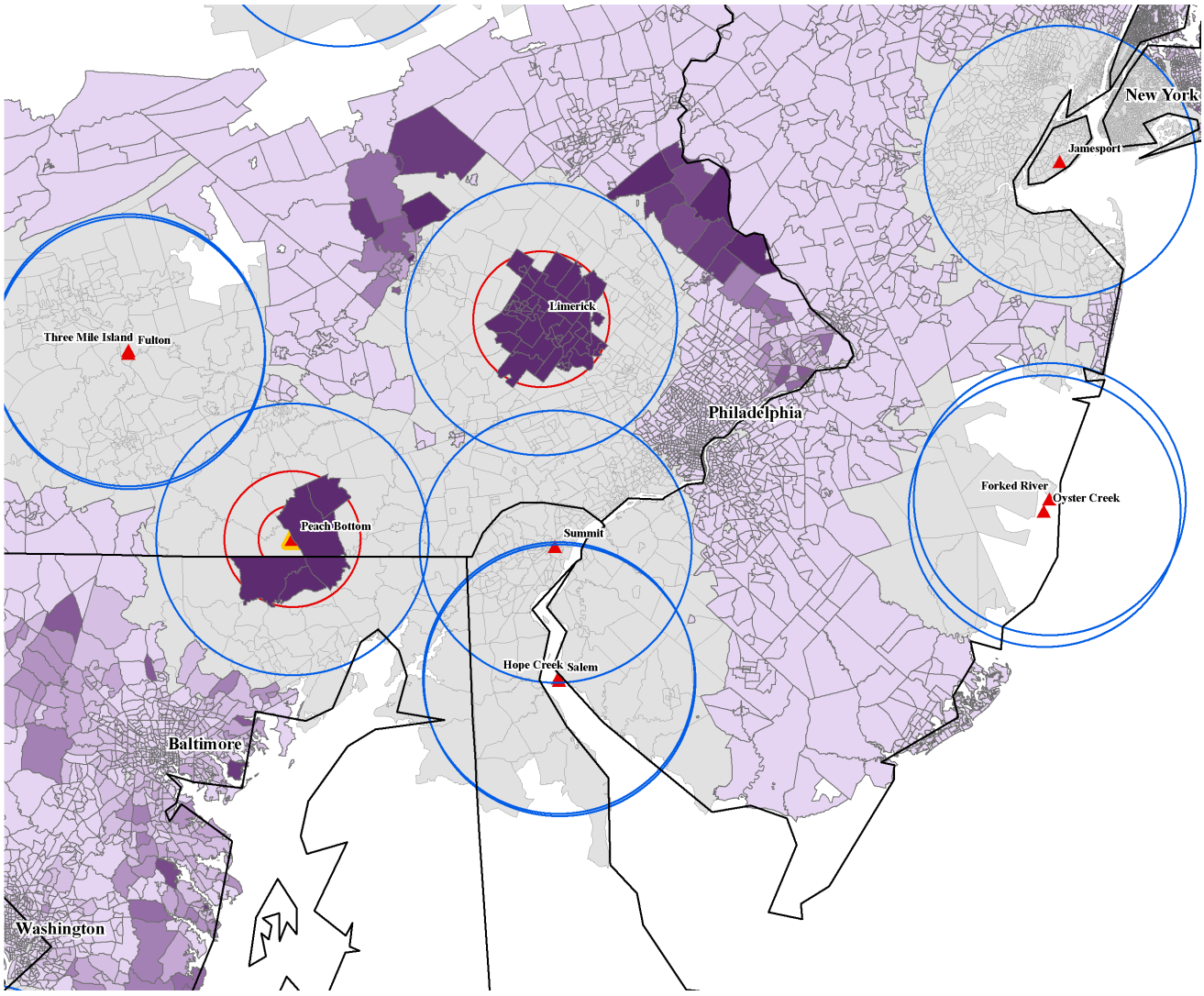


Table 1.7: Summary statistics for plant placement set by 10-mile treatment group and weighted control subsamples.

		Control group subsamples			
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	7.70	5.35	5.38	5.38	5.64*	7.44
Miles to sufficient water source	4.73	5.02	4.93	4.93	4.76	4.73
Pop density (1,000 per sq-mile)	4.76	4.64	4.63	4.62	4.36	3.81
Living in a diff county 5 years ago (%)	24.20	23.76	23.68	23.7	23.92	22.7
Black (%)	16.45	16.35	16.59	16.55	16.3	15.69
Hispanic (%)	4.11	4.71	4.54	4.54	4.71	2.99**
At least 12 years education (%)	27.94	27.51	27.41	27.39	27.76	28.5
Al least bachelors degree (%)	10.07	10.84	10.79	10.78	10.46	11
HH income ($1,000s)	10.07	10.24	10.24	10.24	10.04	10.18
Unemployment rate (%)	3.70	3.73	3.69	3.69	3.68	3.08**
Professional and technical (%)	13.53	13.92	13.91	13.9	13.92	14.68**
Average housing value ($1,000s)	18.41	18.62	18.62	18.59	18.32	21.13**
Vacancy rate (%)	6.50	6.78	6.72	6.74	7.02	7.59
Tenure rate (%)	61.19	60.53	60.62	60.61	60.68	63.38
Multi-units structures (%)	31.83	30.54	30.76	30.81	30.59	27.41
Structures < 5 years old (%)	15.24	16.09	16.05	16.04	16.9	19.94
Structures < 10 years old (%)	28.10	29.62	29.45	29.4	30.62	34.02

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

Table 1.8: Summary statistics for succesfull plant placement set by 10-mile treatment group and weighted control subsamples.

		Control group subsamples			
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	12.79	9.82	9.85	9.93	10.71	10.32
Miles to sufficient water source	2.62	2.75	2.76	2.78	2.64	2.64
Pop density (1,000 per sq-mile)	1.45	1.12	1.16	1.14	1	1.1
Living in a diff county 5 years ago (%)	25.29	22.43	22.27	22.36	21.25	21.61
Black (%)	7.19	8.11	8	7.93	11.01	7.73
Hispanic (%)	1.30	1.25	1.21	1.24	1	0.91
At least 12 years education (%)	32.88	32.03	31.53	31.6	29.57	31.96
Al least bachelors degree (%)	8.97	8.9	9.32	9.37	9.47	8.47
HH income ($1,000s)	10.40	10.25	10.33	10.38	10.13	10.38
Unemployment rate (%)	3.19	3.32	3.19	3.18	3.11	3.35
Professional and technical (%)	14.55	14.19	14.4	14.47	14.17	13.08
Average housing value ($1,000s)	20.41	19.51	19.57	19.64	19.08	18.86
Vacancy rate (%)	5.42	6.31	6.24	6.36	6.62	5.2
Tenure rate (%)	74.15	74.94	74.01	74.25	74.73	75.66
Multi-units structures (%)	14.87	12	12.24	12.16	11.18	12.13
Structures < 5 years old (%)	17.88	18.2	18.01	17.93	19.1	17.97
Structures < 10 years old (%)	30.27	30.59	30.55	30.46	31.74	29.71

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

Table 1.9: Summary statistics for plant closure set by 10-mile treatment group and weighted control subsamples.

		Control group subsamples				
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	9.04	9.59	4.72	4.72	5.75	5.6
Miles to sufficient water source	1.93	1.9	1.93	1.93	2.03	1.89
Pop density (1,000 per sq-mile)	3.42	3.29	5.51	5.51	3.52	3.07
Living in a diff county 5 years ago (%)	20.76	18.05***	22.46	22.46	19.82	19.61
Black (%)	10.76	9.81	9.05	9.05	7.61	5.9
Hispanic (%)	8.48	8.21	18.39	18.39	8.27	7.22
At least 12 years education (%)	32.77	32.79	30.29	30.29	32.98	33.48
Al least bachelors degree (%)	15.54	15.59	14.35	14.35	15.91	16.64
HH income ($1,000s)	36.80	36.85	36.08	36.08	37.81	38.11
Unemployment rate (%)	6.57	6.77	7.69	7.69	6.52	6.3
Professional and technical (%)	15.36	15.41	13.85	13.85	15.49	16.24
Average housing value ($1,000s)	85.08	90.05	85.4	85.4	87.33	88.99
Vacancy rate (%)	7.63	7.67	6.68	6.68	6.5	5.72
Tenure rate (%)	63.71	64.25	58.07	58.07	65.08	66.18
Multi-units structures (%)	32.09	30.92	39.46	39.46	30.28	29.07
Structures < 5 years old (%)	9.29	9.39	6.85	6.85	8.41	9.13
Structures < 10 years old (%)	15.03	15.12	10.86	10.86	13.57	14.8

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

Table 1.10: Summary statistics for Chernobyl accident set by 10-mile treatment group and weighted control subsamples.

		Control group subsamples				
	average within treated	Baseline	BACON	Min/Max	2nd & 98th percentile	4th & 96th percentile
	(1)	(2)	(3)	(4)	(5)	(6)
Miles to populous city centers	13.72	10.64	10.71	10.73	12.06	18.27***
Miles to sufficient water source	2.07	2.07	2.02	2.02	2.06	1.53
Pop density (1,000 per sq-mile)	2.52	2.56	2.67	2.65	2.98	3.15
Living in a diff county 5 years ago (%)	22.95	20.36**	20.36**	20.38**	20.11	21.05
Black (%)	8.93	7.86	7.95	7.91	7.01	2.99***
Hispanic (%)	5.11	5.3	6.21	6.17	5.76	5.11
At least 12 years education (%)	37.92	38.34	37.98	38	38.3	38.26
Al least bachelors degree (%)	16.42	15.13	14.77	14.76	14.57*	14.09
HH income ($1,000s)	21.86	21.83	21.77	21.77	21.67	20.3
Unemployment rate (%)	6.76	7.09	7.13	7.12	7	7.35
Professional and technical (%)	15.79	15.01	14.71	14.71	15	15.66
Average housing value ($1,000s)	55.49	61.75**	54.79	54.79	53.33	50.09
Vacancy rate (%)	6.16	6.01	5.94	5.94	5.81	6.92
Tenure rate (%)	68.27	69.32	68.61	68.63	68.54	64.42
Multi-units structures (%)	29.06	27.37	28.16	28.13	28.4	32.42
Structures < 5 years old (%)	14.16	13.46	12.99	13.02	12.62	11.7
Structures < 10 years old (%)	28.76	27.41	26.49	26.55	25.23	22.06

Notes: *** if significantly different from the treatment group at the 0.001 level. ** at the 0.01 level. * at the 0.05 level. Populous citys are considered those with populations greater than 25,000. Tracts from the large western states without nuclear power plant experience were removed from all control groups.

1.6 Net changes in labor market and demographic mix

1.6.1 Methodology

In order to evaluate the effect of a nuclear power plant event on the labor market and demographic mix, we model our specification based on Davis (2011), which evaluates the impact of the placement of fossil fuel burning power plants. After subsampling the control group and weighting control observations by the relative odds of treatment, the following regression is performed.

$$(D_{i,post-event}) = \gamma N_i + D_{i,controlyear} + \beta X'_{i,controlyear} + \epsilon_i \tag{1.2}$$

,

where $X_{i,controlyear}$ are the tract level covariates listed in Table 2 as well as the square and cubics of household income, the portions of the employed population employed in different industry sectors, and county fixed effects. D is a placeholder for different dependent variables of interest to our investigation. These include log average housing values, vacancy rates, tenure rates, population density, the percent of the population living in a different county five years ago, the unemployment rate, the percent of the employed population employed in professional and technical occupations, the percent of the population with college degrees, and log household income. Least squares estimation of this equation is unbiased if the unobserved determinants of the dependent variable are not correlated with power plant siting decisions after controlling for location fixed effects and $X_{i,controlyear}$ in both the logistic and main regressions. In other words, the errors from the logistic and main regressions must be uncorrelated.

The effect attributable to nuclear treatment (γ) does not control for the changes in demographic, occupational, or housing attribute variables induced by the nuclear event. Therefore, the treatment effect, γ, in a regression with average reported housing value as the dependent variable should include the indirect effects on housing values that feed through

these and other neighborhood characteristics.

1.6.2 Results: Plant placement

We first turn our attention to the effect of treatment when defined around the plant placement set. Recall that these are nuclear power plants placed in the 1970s and whether or not the plants successfully produced commercial power is disregarded. The placement of a commercial nuclear power plant under these conditions causes housing values within the plume exposure pathway (10 mile surrounding area) to decrease by 10.9%[8] (see Table 9). This large net impact on housing values attributable to plant placement includes the effect on home values that feeds through impacts on other important variables predicting housing values. For example, ownership rates drop 3.7 percentage points and the substantial evidence of net outward migration.

The evidence of net outward migration in response to the plants introductions can be seen in both vacancy rates and population density. Vacancy rates within the 10-mile pathway decrease by 1.5 percentage points. Even closer to the plants, within 5 miles, net outward migration is even stronger. Vacancy rates increase by 6.3 percentage points and population density drops by around 930 people per square mile. While net migration may be outward as households move away from new accident risk or the introduction of a local eye sore, there is evidence that other populations were moving in. The evidence of increased availability of high-tech jobs is strongest among the 5-mile treatment area where the percent employed in professional and technical occupations increases by more then 8.5 percentage points. Also in this area, the percentage of the population with a college degree increases by over 5 percentage points.

The regressions with tract average income as the dependent variables show a reduction of over 11.5% in household average income after plant placement. However, as mentioned, this coincides with an increase in the portion of employed individuals in professional and

[8]Relative effect of a dummy variable with estimated coefficient α in a log-linear regression is equal to $e^{\alpha} - 1$ (Halvorsen and Palmquist, 1980).

Table 1.11: Net effect of nuclear power plant placement

.	within 10 miles of plants			within 5 miles of plants	
	BACON	Min/Max	2nd & 98th	Min/Max	2nd & 98th
	(1)	(2)	(3)	(4)	(5)
Log average housing value					
Coef. on N	-0.1153**	-0.1147**	-0.0535	-0.0763	-0.0842
	(0.0476)	(0.0484)	(0.0417)	(0.0539)	(0.0685)
Adj. R2	0.816	0.815	0.793	0.831	0.840
Vacancy rate					
Coef. on N	0.0148*	0.0144*	0.0179**	0.0634***	0.0558***
	(0.0086)	(0.0086)	(0.0078)	(0.0110)	(0.0099)
Adj. R2	0.489	0.491	0.521	0.615	0.641
Tenure rate					
Coef. on N	-0.0387***	-0.0377***	0.0044	0.0443	0.0987*
	(0.0102)	(0.0106)	(0.0313)	(0.0541)	(0.0570)
Adj. R2	0.794	0.794	0.793	0.833	0.844
Population density					
Coef. on N	-0.3856	-0.3695	-0.3512	-0.9288*	-1.1329**
	(0.3322)	(0.3340)	(0.3876)	(0.5128)	(0.5514)
Adj. R2	0.901	0.906	0.885	0.913	0.927
Portion living in a different county five years ago					
Coef. on N	-0.0291	-0.0282	-0.0285	-0.0030	0.0078
	(0.0251)	(0.0255)	(0.0282)	(0.0205)	(0.0233)
Adj. R2	0.659	0.659	0.677	0.718	0.776
Unemployment					
Coef. on N	0.0170*	0.0161*	0.0051	0.0054	-0.0080
	(0.0092)	(0.0091)	(0.0089)	(0.0082)	(0.0091)
Adj. R2	0.700	0.697	0.721	0.696	0.762
Portion in professional and technical occupations					
Coef. on N	0.0076	0.0077	0.0318***	0.0851***	0.0981***
	(0.0089)	(0.0087)	(0.0089)	(0.0156)	(0.0171)
Adj. R2	0.733	0.733	0.708	0.725	0.754
Portion of population with college degrees					
Coef. on N	-0.0174	-0.0173	0.0129	0.0520***	0.0569***
	(0.0107)	(0.0106)	(0.0124)	(0.0169)	(0.0164)
Adj. R2	0.817	0.818	0.779	0.770	0.803
Log average household income					
Coef. on N	-0.1233***	-0.1220***	-0.0429	-0.0752	-0.0155
	(0.0356)	(0.0353)	(0.0308)	(0.0489)	(0.0650)
Adj. R2	0.525	0.526	0.471	0.795	0.796
Num Obs	7994 to 8254	8210 to 8474	4135 to 4257	3500 to 3656	1817 to 1881

Notes: Standard errors clustered by county are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control observations are weighted with the relative odds of treatment estimated using pre-treated variables. All regressions include, in addition to the variables above and those listed in table 2, the square and cubics of log average household income, number of individuals under 18 years old per household, number of individuals over 65 per household, distance to metropolitan centers, and county fixed effects.

technical occupations and higher rates of college degrees. Although nuclear power plants employ highly educated scientists likely earning high wages, the introduction of the power plant may have caused the profile of surrounding businesses to change with migration as well. The type of businesses willing to invest in locations near plants may not themselves employ high wage labor and so average wages are depressed, further lowering the willingness to pay for housing and contributing to the ten percent drop in home values. The fact that the unemployment rate increases by 1.7 percentage points within the plume exposure pathway offers additional support to this conclusion. An alternative explanation for the simultaneous increase in education and drop in median household income could be the outward migration of the original upper or middle tier residents in terms of earnings in response to the introduction of accident risk. The less mobile working class residents may be the greater portion of those to remain.

Next we turn to the results for the successful plant placement set. If we evaluate the net impacts of plant placement in areas where placement resulted in the successfully production of commercial energy, the results change in mostly expected ways (see Table 10). The increases in college educations and the percent employed in professional and technical occupations are even stronger surrounding these successful plants, again with a greater effect in the 5 mile area than in the 10 mile treatment area. More successful operations need more employees and are able to support a larger economy of service workers.

These successful operations reduce the evidence of net outward migration. The impact on population density is insignificant while the percent of the population living in a different county five year ago increases in the 5 mile treatment area by over 34 percentage points. The unemployment rate significantly drops within the 10 and 5 mile treatment areas (0.7 and 4.5 percentage points, respectively). The success of the plant and resulting labor market improvements bolster incomes and housing values, the coefficients on which are now insignificantly negative. Interestingly, while there is evidence of new jobs and inward migration to the 5 mile treatment area, the ownership rates of homes within this ring drops significantly

by over 27 percentage points. This could be evidence that households are less willing to make a long term investment in homes within such close distances to the plant and so are more likely to be affected in the case of a nuclear accident. Another potential explanation is that utility or energy companies could have bought out homeowners to ease the site approval process[9].

1.6.3 Results: Plant closures

In contrast to the results from the plant placement set, plant closures in the 1990s have a net positive impact on housing values. They increase by 10%. Recall that treatment for the plant placement set drops home values by the same magnitude in the 1970s. Also inverting the effects of plant placement, tenure rates and population density increased with the removal of accident risk. However, not all the effects of plant placement are reversed. For example, plant closures further increased the portion of the population with college degrees (14 percentage points), dropped in the unemployment rate (-2 percentage points), and increased household income (19%) within the 5-mile treatment area of these plants. This is possibly due to labor needed for the decommissioning of the nuclear power reactors.

Decommissioning is more than simply closing operations. The Nuclear Regulatory Commission has strict rules governing nuclear power plant decommissioning, which involves the removal of the radioactive fuel but could also include the permanent removal of other major components such as the reactor vessel, steam generators, large piping systems, pumps, and valves. All of which is done under hazardous radioactive conditions requiring highly skilled and expensive labor. The costs of decommissioning the Trojan nuclear plant (a 1130-MWe pressurized-water reactor) was estimated to be $210 million in 1993 dollars, which did not include $42 million for nonradioactive site remediation or $110 million for independent spent fuel storage installation and related fuel management. Relative to other decommissions, the

[9]Consistent data to use as evidence for home purchases by utility companies from existing residents is difficult to find. Records dating back before 1990 are often not available online, and are instead still stored in physical paper or microfiche.

Table 1.12: Net effect of successful nuclear power plant placement

	within 10 miles of plants			within 5 miles of plants		
	BACON	Min/Max	2nd & 98th	BACON	Min/Max	2nd & 98th
	(1)	(2)	(3)	(4)	(5)	(6)
Log average housing value						
Coef. on N	-0.0580	-0.0500	-0.0460	-0.2121	-0.1827	-0.6711
	(0.0547)	(0.0500)	(0.0457)	(0.4276)	(0.4264)	(0.8065)
Adj. R2	0.81	0.809	0.799	0.717	0.724	0.771
Vacancy rate						
Coef. on N	0.0016	0.0021	0.0001	0.0039	0.0033	-0.0264
	(0.0075)	(0.0073)	(0.0093)	(0.0307)	(0.0296)	(0.0339)
Adj. R2	0.484	0.496	0.555	0.871	0.870	0.881
Tenure rate						
Coef. on N	-0.0031	-0.0056	0.0059	-0.2888***	-0.2869***	-0.2693**
	(0.0119)	(0.0116)	(0.0218)	(0.0657)	(0.0633)	(0.1167)
Adj. R2	0.713	0.708	0.672	0.913	0.913	0.911
Population density						
Coef. on N	-0.2338	-0.2076	-0.2489*	0.0148	0.0051	0.1392
	(0.1463)	(0.1457)	(0.1270)	(0.5496)	(0.5435)	(0.9630)
Adj. R2	0.972	0.969	0.975	0.978	0.978	0.985
Portion living in a different county five years ago						
Coef. on N	0.0420**	0.0407*	0.0474	0.3462***	0.3466***	0.4503***
	(0.0199)	(0.0222)	(0.0407)	(0.0985)	(0.0995)	(0.1050)
Adj. R2	0.648	0.651	0.687	0.912	0.912	0.938
Unemployment						
Coef. on N	-0.0074*	-0.0071*	-0.0129**	-0.0455**	-0.0452**	-0.0113
	(0.0038)	(0.0037)	(0.0056)	(0.0200)	(0.0196)	(0.0259)
Adj. R2	0.665	0.653	0.679	0.882	0.881	0.872
Portion in professional and technical occupations						
Coef. on N	0.0238**	0.0235**	0.0205**	0.0581*	0.0613*	0.0983*
	(0.0097)	(0.0101)	(0.0096)	(0.0319)	(0.0314)	(0.0527)
Adj. R2	0.753	0.753	0.742	0.936	0.937	0.959
Portion of population with college degrees						
Coef. on N	0.0250***	0.0264***	0.0088	0.0632**	0.0618**	0.1560***
	(0.0080)	(0.0083)	(0.0104)	(0.0289)	(0.0294)	(0.0368)
Adj. R2	0.824	0.823	0.830	0.934	0.935	0.950
Log average household income						
Coef. on N	-0.0274	-0.0292	-0.0670**	0.7050	0.6878	1.8120*
	(0.0195)	(0.0192)	(0.0267)	(0.7403)	(0.7210)	(0.9364)
Adj. R2	0.741	0.741	0.679	0.493	0.500	0.619
Num Obs	2679 to 2746	3030 to 3103	817 to 839	199 to 265	205 to 276	160 to 191

Notes: Standard errors clustered by county are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control observations are weighted with the relative odds of treatment estimated using pre-treated variables. All regressions include, in addition to the variables above and those listed in table 2, the square and cubics of log average household income, number of individuals under 18 years old per household, number of individuals over 65 per household, distance to metropolitan centers, and county fixed effects.

Trojan plant decommissioning from shutdown in 1993 to license termination in 2002 (beyond the treatment year for our investigation) was considered among the fastest. With these labor demands, it is then not surprising that plant closures in the 1990s increased incomes and dropped the unemployment rate in 2000.

1.6.4 Results: Chernobyl accident

In response to the Chernobyl accident, population density around operating plants significantly increased by approximately 300 people per square mile between 1980 (pre-accident) and 1990. In addition, the percent of the population with college degrees (1.4 percentage points within 10 miles and 6.1 percentage points within 5 miles) and the portion employed in professional and technical occupations (3 percentage points within 10 miles and over 4 points within 5 miles) also increased. After the Three Mile Island and Chernobyl accidents, the safety regulations on commercial nuclear power reactors were heightened requiring more engineers and inspectors to maintain safe and legal operations. The net effects of the Chernobyl accident on population density, education levels, and the portion of the employed population in professional and technical occupations reflect this. However, the labor market improvements were unable to support housing values. The net impact on housing values in response to the accident was another negative shock, dropping by 7.3% within the plume exposure pathway.

Table 1.13: Net effect of nuclear power plant closures

.	within 10 miles of plants		within 5 miles of plants	
	Min/Max	2nd & 98th	Min/Max	2nd & 98th
	(1)	(2)	(3)	(3)
Log average housing value				
Coef. on N	0.1121*	0.0840	0.3742***	0.3374***
	(0.0672)	(0.0707)	(0.1073)	(0.0928)
Adj. R2	0.915	0.941	0.894	0.916
Vacancy rate				
Coef. on N	0.0180	0.0084	0.0078	0.0040
	(0.0125)	(0.0102)	(0.0118)	(0.0126)
Adj. R2	0.907	0.923	0.981	0.979
Tenure rate				
Coef. on N	0.0554***	0.0649***	0.0821	0.0398
	(0.0198)	(0.0166)	(0.0576)	(0.0577)
Adj. R2	0.909	0.897	0.820	0.816
Population density				
Coef. on N	0.3445***	0.3461*	0.5619***	0.4842***
	(0.1192)	(0.1976)	(0.1156)	(0.1650)
Adj. R2	0.991	0.981	0.993	0.993
Portion living in a different county five years ago				
Coef. on N	0.0276	0.0326**	0.0148	0.0157
	(0.0187)	(0.0164)	(0.0357)	(0.0404)
Adj. R2	0.691	0.628	0.734	0.735
Unemployment				
Coef. on N	0.0126	-0.0006	-0.0225*	-0.0202*
	(0.0099)	(0.0077)	(0.0122)	(0.0103)
Adj. R2	0.497	0.524	0.707	0.702
Portion in professional and technical occupations				
Coef. on N	0.0010	-0.0023	0.0047	0.0026
	(0.0130)	(0.0141)	(0.0306)	(0.0206)
Adj. R2	0.792	0.782	0.667	0.684
Portion of population with college degrees				
Coef. on N	0.0597	0.0482	0.1575***	0.1432***
	(0.0450)	(0.0437)	(0.0355)	(0.0385)
Adj. R2	0.847	0.861	0.856	0.862
Log average household income				
Coef. on N	-0.0282	-0.0143	0.2400***	0.1739***
	(0.0666)	(0.0638)	(0.0681)	(0.0585)
Adj. R2	0.879	0.898	0.859	0.879
Num Obs	2861	1793	310 to 554	290 to 487

Notes: Standard errors clustered by county are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control observations are weighted with the relative odds of treatment estimated using pre-treated variables. All regressions include, in addition to the variables above and those listed in table 2, the square and cubics of log average household income, number of individuals under 18 years old per household, number of individuals over 65 per household, distance to metropolitan centers, and county fixed effects.

Table 1.14: Net effect of the Chernobyl accident

.	within 10 miles of plants			within 5 miles of plants	
	BACON	Min/Max	2nd & 98th	Min/Max	2nd & 98th
	(1)	(2)	(3)	(4)	(5)
Log average housing value					
Coef. on N	-0.0746**	-0.0721**	-0.0596	-0.0496	0.2114***
	(0.0334)	(0.0323)	(0.0387)	(0.0624)	(0.0729)
Adj. R2	0.960	0.960	0.963	0.980	0.994
Vacancy rate					
Coef. on N	-0.0040	-0.0032	0.0087	-0.0484*	-0.0360
	(0.0127)	(0.0126)	(0.0140)	(0.0286)	(0.0367)
Adj. R2	0.708	0.709	0.747	0.894	0.938
Tenure rate					
Coef. on N	-0.0460**	-0.0449***	0.0150	0.0983	-0.0534
	(0.0178)	(0.0172)	(0.0155)	(0.0599)	(0.1425)
Adj. R2	0.840	0.840	0.853	0.785	0.891
Population density					
Coef. on N	0.2954*	0.2961*	-0.0191	0.2993**	0.3189
	(0.1506)	(0.1518)	(0.1120)	(0.1298)	(0.2625)
Adj. R2	0.986	0.986	0.991	0.998	1.000
Portion living in a different county five years ago					
Coef. on N	0.0071	0.0072	-0.0382**	-0.0772**	0.0630
	(0.0129)	(0.0128)	(0.0169)	(0.0312)	(0.0708)
Adj. R2	0.639	0.641	0.657	0.819	0.898
Unemployment					
Coef. on N	-0.0116	-0.0118	0.0034	-0.0123	-0.0253
	(0.0087)	(0.0088)	(0.0063)	(0.0101)	(0.0216)
Adj. R2	0.698	0.696	0.726	0.862	0.853
Portion in professional and technical occupations					
Coef. on N	0.0315***	0.0313***	0.0243***	0.0546***	0.0448
	(0.0089)	(0.0091)	(0.0088)	(0.0174)	(0.0304)
Adj. R2	0.824	0.823	0.826	0.907	0.952
Portion of population with college degrees					
Coef. on N	0.0146**	0.0152**	-0.0062	0.0589**	0.1220***
	(0.0074)	(0.0074)	(0.0124)	(0.0258)	(0.0392)
Adj. R2	0.874	0.873	0.857	0.902	0.903
Log average household income					
Coef. on N	-0.0100	-0.0108	0.0503	0.1202	0.0923
	(0.0233)	(0.0228)	(0.0317)	(0.0756)	(0.1361)
Adj. R2	0.893	0.892	0.888	0.843	0.796
Num Obs	8548	8713 to 8718	3125 to 3126	1300	516

Notes: Standard errors clustered by county are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control observations are weighted with the relative odds of treatment estimated using pre-treated variables. All regressions include, in addition to the variables above and those listed in table 2, the square and cubics of log average household income, number of individuals under 18 years old per household, number of individuals over 65 per household, distance to metropolitan centers, and county fixed effects.

1.7 Amenity value of nuclear power plants

1.7.1 Methodology

After a major nuclear event, we expect households to respond by sorting themselves along a distance gradient according to their preferences. These preferences may be for greater distance from the plant to avoid the perceived risk of a nuclear power plant or for closer access to the labor market improvements we discussed in the previous section. We capture these changing labor market improvements and control for their impact on housing demand to isolate the amenity value of nuclear power plants[10]. In the case of the Chernobyl Accident treatment, this model specification estimates the change in the amenity value induced by the new information on potential damages. We do this by including both pre- and post-treatment values of the same covariates used in the previous regressions, which describe different features of the housing and labor market.

Treating each tract in a given year as a single observation, the general model is

$$\ln(P_{it}) = \rho_1 N_i + \rho_2 T_t + \gamma N_i \times T_t + \beta X'_{it} + \epsilon_{it} \tag{1.3}$$

where P_{it} is the log average reported housing value in tract i in period t, $T_t = 1$ in the year following the event and 0 in the control year. N_i is our nuclear proximity treatment variable. These dummy variables control for the time invariant differences in these regions. T_t and N_i absorb contemporaneous non-nuclear related shocks in both groups and the time invariant differences between groups, respectively. X_i includes the same tract level covariates as those used in the previous regression specifications to evaluate the net effects of nuclear power plant events. However, for these DID regressions, both control year and treatment year values are used.

The treatment effect and coefficient of interest is γ. This is the difference in home values between the treatment period and control period in the treatment group exceeding the same

[10] See Appendix C for a discussion of the hedonic method applied to the context of nuclear power proximity.

difference observed in the control group. This is why the regression structure is referred to as a Difference-in-Difference approach.

The inclusion of the covariates X controls for time varying differences in treatment and control groups by including variables that capture the demographic mix and labor market features of tracts after the nuclear event. This reflects the evolving nature of neighborhoods and their populations over time. In this way we are also eliminating the indirect effect of the nuclear event on housing values. These indirect effects act on housing values through these changing demographics and labor market features. For example, historically home values tend to increase with income. Higher incomes increase the reservation price for housing. In the case of nuclear events, some original residents migrate away at the same time that technical occupations increase, pulling in more educated populations. This puts both upward and downward pressure on housing values. The inclusion of post-event employment variables will separate these effects from the estimated value of γ. The impact on housing values that remain should be the discount associated with proximity to nuclear power plants. This reveals the disamenity value of the dichotomous treatment, close or not close to nuclear power associated with the nuclear event.

In order for the DID effect to reflect the effect of the nuclear event, we must expect both our treatment and control groups to be subject to the same non-nuclear related housing market shocks and that their prices are expected to evolve similarly over time in the absence of the nuclear intervention **?**. The control group subsampling and weighting system, which heavily weights the more appropriate control observations, play a major role in ensuring this is the case, as does the use of county-year fixed effects.

1.7.2 Results: Plant placement

The results from the successful plant placement set, those plants placed during the 1970s that continued on to commercial operation, value the disamenity associated with nuclear power plant placement between 11.6 and 47.1% of home values (see Table 13). The results

from the plant placement set, the full set of nuclear power plants placed in the 1970s both successful and not, yield insignificant results regardless of the subsample and treatment definition used. It may be the case that unsuccessful placements met their demise after long public battles, which reduced homeowners expectation that a nuclear power plant would actually be in the vicinity in the foreseeable future.

Table 1.15: The amenity/disamenity value of successful plant placement

.	Baseline (1)	BACON (2)	Min/Max (3)	2nd & 98th (4)	4th & 96th (5)
Within plume exposure pathway (10 miles) of plants					
Coef on N*T	-0.0907***	-0.0495	-0.0449	-0.0010	-0.0389
	(0.0218)	(0.0307)	(0.0337)	(0.0510)	(0.0485)
Coef on T	1.2716***	0.7022***	0.7177***	0.6986***	0.8004***
	(0.1506)	(0.1234)	(0.1225)	(0.1833)	(0.1368)
Coef on N	-0.1257***	-0.1400**	-0.1388**	-0.0924	-0.2021***
	(0.0308)	(0.0565)	(0.0617)	(0.1007)	(0.0571)
Obs	14090	5412	6120	1650	897
Adj. R2	0.951	0.959	0.959	0.959	0.967
Within 5 miles of plants					
Coef on N*T	-0.1235***	-0.6092**	-0.6366**	-0.6237**	1.4265**
	(0.0277)	(0.2491)	(0.2471)	(0.2535)	(0.5789)
Coef on T	1.4955***	2.3039***	2.3408***	2.5528***	1.6658**
	(0.2732)	(0.6495)	(0.6593)	(0.8187)	(0.6569)
Coef on N	-0.0951*	-0.6947**	-0.7029**	-0.7887	-0.6838
	(0.0549)	(0.3174)	(0.3275)	(0.5695)	(0.5157)
Obs	13963	428	442	333	123
Adj. R2	0.942	0.972	0.972	0.971	0.988

Notes: Standard errors clustered by county are in parentheses; *** p¡0.01, ** p¡0.05, * p¡0.1. Control observations are weighted with the relative odds of treatment. All regressions include, in addition to the variables above and those listed in table 1, the square and cubics of household income, number of individuals under 18 years old per household, number of individuals over 65 per household, distance to metropolitan centers, and county by year fixed effects.

1.7.3 Results: Plant closures

The results from the plant closure set, those plants that closed in the 1990s, reinforce the results from successful plant placement and the hypothesis that nuclear power plants introduce negative externalities to the local community. The amenity value of plant closures is positive and significant in the plume exposure pathway, with larger effects in the 5 mile areas surrounding closed plants (see Table 14).

Table 1.16: The amenity/disamenity value of plant closures

	Baseline (1)	Min/Max (2)	2nd & 98th (3)	4th & 96th (4)
Within plume exposure pathway (10 miles) of plants				
Coef on N*T	0.2459**	0.2218***	0.1999***	0.7756***
	(0.1015)	(0.0597)	(0.0590)	(0.2103)
Coef on T	0.0693	0.0302	-0.0673	-0.6059***
	(0.0872)	(0.1369)	(0.1631)	(0.2145)
Coef on N	-0.2671*	-0.1499***	-0.1229***	-0.0090
	(0.1522)	(0.0289)	(0.0303)	(0.0353)
Obs	18545	5722	3586	1720
Adj. R2	0.894	0.937	0.936	0.946
Within 5 miles of plants				
Coef on N*T	0.3263***	0.1670**	0.1164	0.2219*
	(0.0838)	(0.0795)	(0.1051)	(0.1311)
Coef on T	0.0940	-0.5831***	-0.5297***	-0.5367***
	(0.0868)	(0.1605)	(0.1533)	(0.0804)
Coef on N	-0.3314**	-0.1445***	-0.1409***	-0.1499***
	(0.1560)	(0.0347)	(0.0377)	(0.0333)
Obs	18415	678	630	498
Adj. R2	0.884	0.963	0.962	0.954

Notes: Standard errors clustered by county are in parentheses; *** p¡0.01, ** p¡0.05, * p¡0.1. Control observations are weighted with the relative odds of treatment. All regressions include, in addition to the variables above and those listed in table 1, the square and cubics of household income, number of individuals under 18 years old per household, number of individuals over 65 per household, distance to metropolitan centers, and county by year fixed effects.

1.7.4 Results: Chernobyl accident

The results from the accident set, those operating commercially before 1980 and continuing through 1990, so no significant impact on the amenity values associated with nuclear power plants. A potential explanation for this is the preexisting sorting according to preferences that would have taken place after this set of plants was initially placed. At the time of the accident, those with accident risk aversion had already moved away from the area.

1.8 Conclusion

Although nuclear power plants introduce high-tech jobs and many labor market improvements such as lower unemployment rates and a higher level of education, they cause housing values to decrease overall. Homeownership rates in local areas are also lower after the introduction of nuclear power plants and after the Chernobyl accident demonstrated the extent of damage possible. Homeowners experience negative externalities from living near nuclear power plants evidenced by lower relative housing prices while controlling for labor market improvements and net outward migration. The effect of plant closures consistently reverses the effect of placement, with the exception of continued increases to education levels and technical occupations held by those living in very close proximity to the plant. This is likely evidence of high skills needed for decommissioning.

The recent nuclear power plant closures in 2013 and the expected closure in 2014, as well as the Fukushima accident in 2010, offer an opportunity for further research as a natural extension to this paper. National datasets of local home values are far more available in recent years. For example, Zillow, an online home-related marketplace, publishes monthly home value time series data going back to 1997 for a majority of zip codes in the US. The decennial census, by only collecting enough households for a national snapshot at the census tract level every ten years is vulnerable to other events besides nuclear placement, closures, or accidents that can occur during the relevant decades. The greater frequency of Zillow data avoids these issues, though is less granular.

Chapter 2

HETEROGENEOUS RESPONSES TO SOCIAL NORMS FOR WATER CONSERVATION

2.1 Introduction

Water utilities, particularly those in the Western United states, face pressure to meet rising water demand. California, for example, has required that utilities reduce demand by 20% by 2020[1] and Texas recently adopted a $2 billion water infrastructure fund.[2] To achieve these conservation goals, water utilities are exploring every strategy in their arsenal. Incorporating social comparisons into information campaigns is an attractive policy instrument aimed at changing consumer behavior when the price mechanism is not feasible due to market structure or political concerns. Social comparisons are particularly well-suited for markets where consumers may have difficulty optimizing their level of use due to a lack of salience over personal consumption quantities, the costs of water or energy saving investments, or the commodity's price signal itself.

Relative information has the additional advantage of leveraging motivations that have been extensively studied by psychologists and behavioral economists. This fits into a broader context of applying behavioral economics in public policy by making relatively small, and seemingly innocuous, changes to a decision framework such as changing the default option. Numerous examples of behavioral economics in policy initiatives span food choice, retirement allocation, organ donation, and utility demand. (Downs et al., 2009; Johnson and Goldstein, 2003; Benartzi and Thaler, 2007; Ferraro et al., 2011; Dolan and Metcalfe, 2011). Opower, an early pioneer for social comparisons in utility demand, sends households information comparing their energy consumption to a group of their peers. Opower's proven success, with almost fifty utility clients spanning six countries and saving over 2.5 billion kilowatt hours, is a shining example of the new application of behavioral economics to achieve environmental goals.[3] Research on social norms for water consumption assesses how this relatively new tool fits into the array of conservation options at the disposal of water managers seeking to

[1]See http://www.swrcb.ca.gov/water_issues/hot_topics/20x2020/ for more information.

[2]See http://openstates.org/tx/bills/83/SJR1/ for the full text of the bill.

[3]Information on Opower can be found at their website - http://opower.com/customers.

influence demand with a very different profile than energy.

We harness a rich dataset from WaterSmart Software, a clean technology company establishing itself as the Opower of water, to analyze the effects of social norms as a water conservation tool. WaterSmart is currently working with twelve utilities, in different stages of development. Three began as randomized pilot programs and are mature enough for analysis. Randomized field experiments are an exciting area of economic research that addresses issues of endogeneity due to selection into treatment in the data generating process; see among others (List and Lucking-Reiley, 2000; Harrison and List, 2004; Gneezy and List, 2006; Landry et al., 2006; Levitt and List, 2009). The randomization process allows for simple estimation of average treatment effects with easily defensible assumptions. This contrasts to using observational data that requires more stringent assumptions and complicated econometric techniques to identify causal effects. Eliminating endogeneity allows us to devote energy to understanding the mechanisms through which consumers reduce consumption by analyzing the heterogeneity in treatment response.

Exploring heterogeneity uncovers the mechanisms through which the program reduces consumption as well as identifies ways to improve the program's performance and cost effectiveness. Our results show that there is significant heterogeneity in the response to the program across initial water use within pilots and stark differences in the average treatment effect (ATE) between pilots. The pilot with the largest ATE, at 5.2%, saved three times as much water as the pilot with the smallest. We relate the pattern in ATE magnitudes to general pilot differences in Section 2.5. Within-pilot heterogeneity shows that baseline water consumption, defined as average water use prior to the treatment, has the most significant and consistent effect on treatment response. Based on quantile regressions and regressions that interact the treatment effect with baseline consumption, households that use more water have a greater response to treatment in all three pilots. Households that use high levels of water have a greater financial incentive to reduce consumption and may be more able to significantly reduce consumption through low marginal cost mechanisms such as fixing leaks

and changing irrigation practices. Other conservation programs using social norms found a similar pattern of results for both energy (Allcott, 2011b) and water (Ferraro and Miranda, 2013).

The intra-pilot heterogeneity of baseline consumption is not reflected in the results for ATEs across pilots. Though there are only three pilots to analyze, the magnitude of the ATE is inversely related to the average pilot-level consumption. The incongruence between the inter- and intra-pilot effects of baseline water use is not surprising and lies in the nature of the social norm campaign. The social norm message compares a household's consumption to a group of peers with similar determinants of water demand in the same utility. For households with similar, or lower, consumption than their local peers the program may cause little, or negative, water savings due to the favorable normative message they receive.[4] By design, the social norm message does not influence consumer behavior through the relative water use between pilots. To explain inter-pilot heterogeneity, differences in specific utility policies and the demographics of the customer base are more relevant and will be discussed later.

While baseline water use influences the intra-pilot heterogeneity in treatment response in a consistent way across pilots, this is not the case for ideology and housing values, which we use as a proxy for income. The significant results for conditional average treatment effects (CATEs) from interactions with ideology quintiles display a different role for within-pilot variation in ideology in the separate pilots. Interactions with housing value quintiles only produce significant results for the CATE in one of the three pilots.

We examine the mechanisms through which consumers reduce water use in response to the social norm by examining the time pattern of savings and interactions with existing conservation programs. Since monthly savings do not increase in the summer months the behavioral response likely focuses on indoor water use. Different patterns across utilities imply that in some pilots consumers make investments in consumer durables or establish

[4]In other studies (Allcott, 2011b; Ferraro and Miranda, 2013) the categorical specification of the normative message did not cause low-use household to increase consumption. In future research we hope to analyze the effect of the specific messages through a regression discontinuity design similar to Allcott (2011b).

permanent behavioral changes, whereas in one pilot the savings appear to be transient even while messaging persists. Interacting the social norm with existing conservation programs finds that social norms act as a complement to existing programs. The social norm message contributes to higher participation rates in other utility programs, which then contributes up to 20% of the average treatment effect assigned to the social norm message.

This paper is organized as follows. The next section briefly discusses the motivations for the use of information campaigns and social norm messaging over pecuniary mechanisms. The following section presents the existing literature estimating the impact of social norm campaigns in the context of utility use. Then, Section 2.4 discusses the background of WaterSmart, describes the data, and provides details on the experimental design. Section 2.5 overviews the estimation strategy and presents the empirical results. Section 2.6 provides plans for future research as we receive more data. Concluding remarks are offered in Section 2.7.

2.2 Theoretical Background

Using social norms in public policy is a relatively recent development. Although a water utility's tariff structure is an important tool in managing demand, there are a number of reasons why the study of non-pecuniary approaches is appealing. First, although some utilities incorporate scarcity pricing into water rates the primary objective in designing rate structures remains cost recovery. Those utilities that do account for scarcity may not fully price the complex risks to water supplies associated with climate change. Second, even under the assumption that utilities appropriately incorporate environmental externalities and increasing scarcity into price schedules, many utilities face zero profit constraints that limit their ability to raise rates.[5] In addition to zero profit constraints, the public utilities that serve most major population centers may find it difficult to raise rates due to concerns

[5]While utilities could restructure water rates to raise the marginal price while leaving the average price unchanged, there is a debate whether consumers actually respond to the average or marginal price of water (Ito, 2013; Nataraj and Hanemann, 2011; Gaudin, 2006; Shin, 1985).

over equity and political backlash by constituents of local governments. While the price mechanism is the most efficient way to reduce consumption, other instruments must be considered when raising prices or introducing Pigouvian taxes is not feasible. Command and control interventions, such as imposing mandatory restrictions on outdoor water use, are a common alternative to raising water rates but are shown to reduce consumer welfare (Grafton and Ward, 2008; Mansur and Olmstead, 2012). In this policy climate water managers must seek alternative mechanisms to reduce demand .

Incorporating social comparisons into information campaigns is an attractive policy instrument aimed at changing consumer behavior when the price mechanism is not feasible due to market structure or political concerns. These benefits are augmented in settings where it is difficult for consumers to optimize their level of use due to a lack of salience over personal consumption quantities, the costs of water or energy saving investments, or the commodity's price signal itself. There are reasons to suspect that consumers do not, or are unable to, choose their monthly water use based on a careful maximization process. Given that the outlays on water generally comprise less than 2% of household income, and that a proportion of water use will remain constant to fill basic human needs, many consumers may not devote much attention to their water bill. The consumers that do scrutinize their bills not only need to know the water intensity of various household activities, but also how to translate complex rate structures into marginal costs. The complexity in water rates may cause consumers to respond to average price instead of marginal price as suggested by Ito (2013, 2012); leading to increased water use since the average price can be substantially lower than marginal price in utilities with increasing block rates. Sexton (2010) shows that automatic energy bill payments increase energy consumption by decreasing the salience of the price signal and Gaudin (2006) finds that water utilities that do not present the marginal price on their bills have lower price elasticity. If households are consuming at the optimal level prior to receiving a social comparison, the information should have no effect on their water use. These studies, and the results of this paper, suggest that salience matters, and information campaigns in

water may work in part by bringing water use to the forefront of consumer decision.

Relative information has the additional advantage of leveraging motivations that have been extensively studied by psychologists and behavioral economists. Elster (1989) attempts to reconcile the divide between theories in economics, where rational agents are "pulled" by incentives, and sociology, where agents are "pushed" by their surrounding environment. Thaler and Sunstein (2003) describe a third route between these two extremes whereby agents still make rational decisions, but can be "nudged" by changing seemingly inconsequential elements in the choice environment. Comparing household use to a group of peers establishes a baseline level of conventional consumption. Knowledge of this convention may generate civic pressure to refrain from resource profligacy. Alternatively, discovering that personal use levels are far above the convention may signal the availability of low cost investments to reduce use. Recent research applies these ideas from behavioral economics to a range of public policy objectives spanning food (Downs et al., 2009), organ donation (Johnson and Goldstein, 2003), retirement savings (Benartzi and Thaler, 2007), and several examples within environmental economics (Allcott, 2011b; Ferraro et al., 2011; Dolan and Metcalfe, 2011). Research on social norms for water consumption assesses how this relatively new tool fits into the array of conservation options at the disposal of water managers.

2.3 Existing Literature

A series of papers by Hunt Allcott and coauthors (Allcott and Mullainathan, 2010; Allcott, 2011a,b; Allcott and Rogers, 2012b,a; Allcott and Greenstone, 2012) on various dimensions of energy economics help establish a role and set of techniques for applied environmental economists to investigate motivations explored in behavioral economics. Allcott's work analyzing data from Opower (Allcott, 2011b; Allcott and Rogers, 2012a,b) parallels our research on water conservation. Allcott (2011b) finds that Opower's program of providing social comparisons along with technical information reduced energy demand by an average of 2%. There is considerable heterogeneity in the treatment effect across the consumption

distribution, with the right tail experiencing the largest savings.

Allcott's work on Opower serves as a template for the analysis since water and energy share many common features. However, there are key differences between water and energy that warrant studying the efficacy of social norms in the water sector. First, while many regions experience seasonal peak demand in the summer for both energy and water, the marginal value of activities driving the peak is likely different. Air conditioning, a major component of peak energy demand, has significant health benefits in combating heat stroke (Semenza et al., 1996). Conversely, research shows that summer demand due to discretionary outdoor use has higher demand elasticity than indoor use (Mansur and Olmstead, 2012; Worthington and Hoffman, 2008; Dalhuisen et al., 2003; Espey et al., 1997) suggesting a lower marginal value for activities shaping peak demand. Additionally, while the challenge of peak energy demand is the cost of supplying daily maximum demand with the most expensive energy production, water utilities are primarily concerned with seasonal demand peaks that are correlated with stressed summer water supplies. Second, energy externalities are primarily global in nature associated with climate change risks from greenhouse gas emissions;[6] conversely, water consumption externalities such as groundwater exhaustion and reducing in-stream flows for ecosystem services have predominantly local impacts. Consumers may be more receptive to limiting externalities whose effects are concentrated within their region as opposed to providing the global public good of emission reductions. Due to both the consumption patterns and localism of environmental impacts, a priori, we may expect to see a different pattern of response to social norms in energy and water.

Ferraro, Price, and Miranda (Ferraro et al., 2011; Ferraro and Miranda, 2013; Ferraro and Price, 2013) analyze an information campaign using social norms in Georgia to reduce water consumption during a severe drought. Ferraro et al. (2011) explicitly test the efficacy of different combinations of technical information, an appeal to support a public good, and a social comparison and find that a social comparison along with technical information is most

[6]There are local health benefits from energy conservation associated with air quality depending on the fuel source and location of power generation.

effective in reducing consumption. Investigations of the heterogeneity for the same program find that wealthy consumers with high water use are most responsive (Ferraro and Miranda, 2013) and that the effect decreases over time (Ferraro and Price, 2013). A major difference between that program and our study area is that WaterSmart applies a continuous treatment over many months, whereas the Georgia program had only one mailing and one follow-up. The dynamic effects of the program in Georgia are analyzed after the program is over, while in our setting we investigate whether conservation gains from social norms are sustainable over many treatment periods. Additionally, that project only encompassed one service area; so while intra-program heterogeneity can be explored it is not possible to compare responses across various regions. Lastly, the program was in response to a severe drought during which public perception towards water consumption was likely heightened.

In assessing the efficacy of social norms as a water conservation tool it is essential to place this instrument in the context of multiple concurrent policies, generally defined as demand side management. Utilities provide home water audits to provide a personal recommendation for water savings as well as offer rebates to replace toilets and washing machines and convert lawns to low-water vegetation. Part of the attraction of social norms is that existing techniques are either not cost effective or do not produce the necessary reductions in demand. Bennear et al. (2013) shows that rebates for high-efficiency toilets do very little to incentivize new consumers; rather the policy merely rewards or speeds up purchases that would naturally occur in absence of the rebate. Mansur and Olmstead (2012) and Grafton and Ward (2008) find that outdoor watering restrictions implemented in arid regions or during extreme droughts cost households over $100 per irrigation season. Past research finds that non-pecuniary programs are less effective than price and mandatory restrictions (Renwick and Archibald, 1998) and that more research is needed to explore policy interactions (Renwick and Green, 2000). Section 2.4.3 goes into details on the existing conservation programs in our study area and Section 2.5.4 conducts an empirical analysis of the interplay of social norms with these programs.

2.4 Background & Data

2.4.1 WaterSmart Software

WaterSmart Software[7] (henceforth WaterSmart) is a clean technology company that contracts with water utilities to reduce demand by providing consumers with Home Water Reports (HWR) containing a social comparison of water use and technical advice. Additionally, WaterSmart assists utility companies with the analysis and interpretation of water metering data. Utility companies desire reductions in demand to offset costly temporary water purchases, delay investments in new supplies, and meet regional or state objectives. Raising rates is often politically infeasible, and incentive programs such as toilet rebates are expensive and may not meet goals of additionality (Bennear et al., 2013).

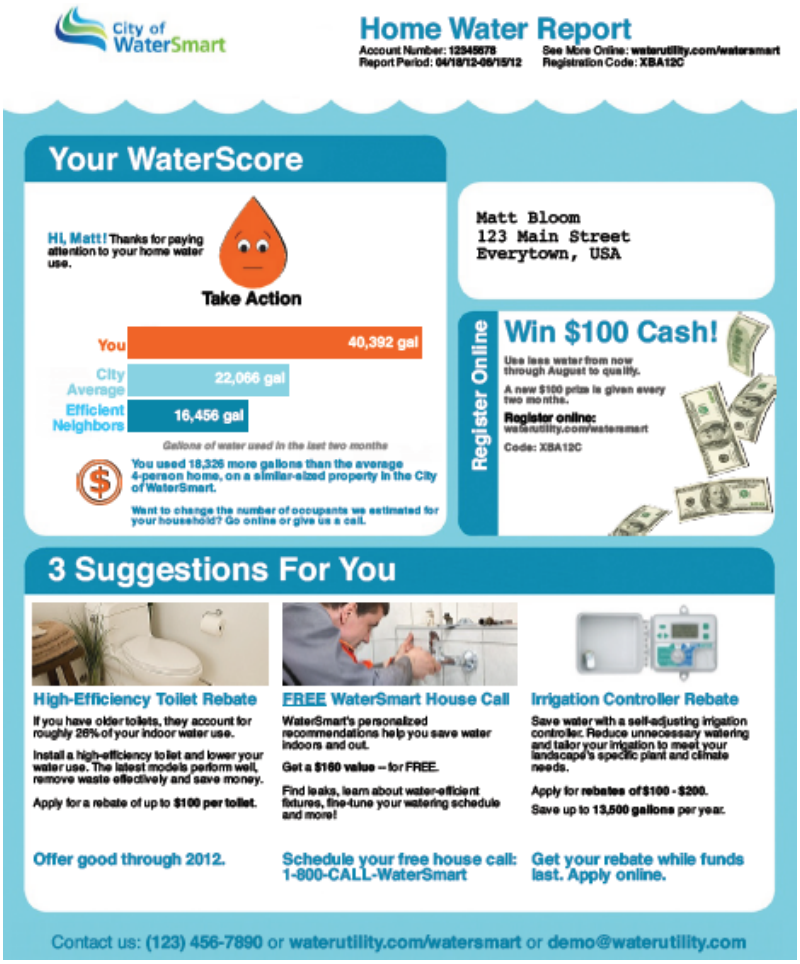
WaterSmart's business model is to send customized Home Water Reports (HWR) with three sections to customers. Some HWRs are mailed to households, while others are sent via email. The main component is a social comparison that contrasts a household's water consumption to that of its peers and then provides a three-tiered qualitative WaterScore based on their relative position. The second section is a list of three personalized recommendations for strategies to save water and rebates available from the utility. Lastly, there is a space that offers incentives for signing up for an online account for more detailed data on water use and additional water conservation tips. Figure 2.1 shows an example of the HWR for an above-average home[8]. A sample of the additional information available through the web portal is presented in Figure 2.2.

---

[7]Additional information on WaterSmart is available on their website http://www.watersmartsoftware.com/ .

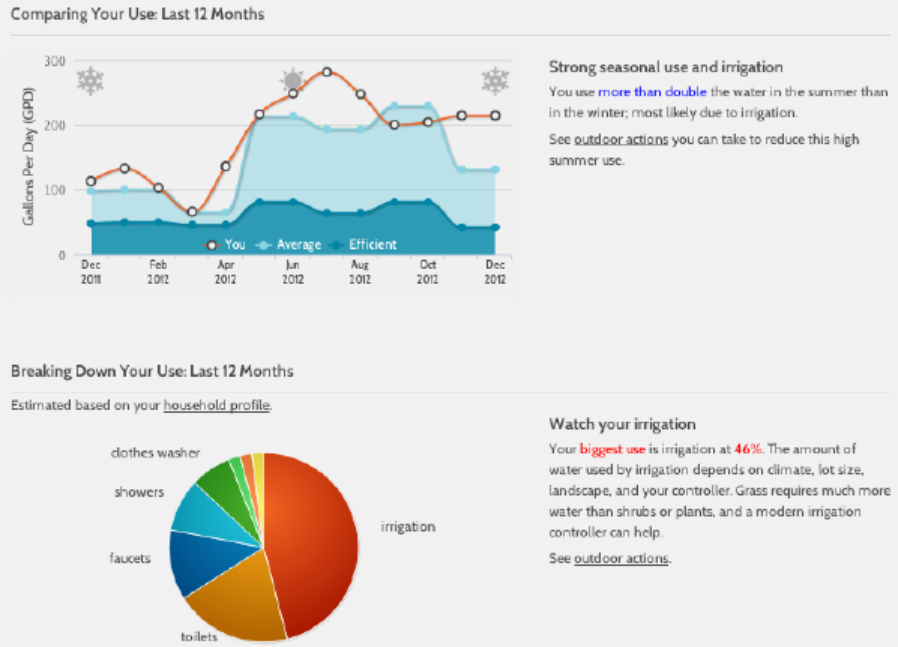
[8]The other categories are "Good" and "Wise" and are very similar to this HWR and are available upon request.

Figure 2.1: Home Water Report



Notes: Home Water Reports have WaterScores of "Wise", "Good", or "Take Action". The bottom panels contain suggestions that are customized based on household data. The right "Win $100" offers different incentives or messages

Figure 2.2: Online Web Portal



Notes: Online web portal is available to the treatment group and can be viewed as an additional "opt-in" treatment over the base treatment. It offers additional data on consumption, utility rebates, and technical information to help reduce water consumption.

2.4.2 Data

We combine several sources of data for the analysis. WaterSmart provides the water metering records at the household level, as well as structural features of the house and treatment status from three California utilities in pilot programs. Daily weather data is obtained from the National Oceanic and Atmospheric Administrations National Climatic Data Center.[9] We match households to the nearest weather station with available temperature and precipitation data and aggregate daily values within the irregular read periods of household water read observations.

Income data from the Census Bureau is not available at a fine enough geography to proxy for household level income. Instead, we use block level median housing values from Zillow, an on-line real estate company with a database of millions of homes and sales transactions across the US.[10] Lastly, we gather voting records from the Statewide Database hosted by the University of California, Berkeley to create ideological indices capturing the environmental sympathies of individual block groups.[11] The Data Appendix (coming soon) details the process of creating a working dataset through all the primary data sources, as well as quality assurance measures taken before conducting the analysis.

We expect that environmental attitudes influence the manner in which households interpret the social norm for water conservation. The Green Ideology Index (GII) is formed using six census block level measures from the 2008 and 2010 California elections: the percent who voted Democrat for U.S. Senator in 2008 and Governor in 2010, the percent who voted yes on "pro-environment" Propositions 7 (2008), 10 (2008), and 21 (2010), and the percent who voted no on Proposition 23 (2010), which would have suspended the state's 2006 Global Warming Solutions Act.[12] For each measure, we assign a score of 100 to the top

[9]We use the temperature and precipitation data from the Global Historical Climatology Network-Daily (GHCN-D), which is publicly available to download at http://www.ncdc.noaa.gov/data-access/quicklinks#ghcn.

[10]http://www.zillowblog.com/research/2012/01/21/zillow-home-value-index-methodology/

[11]The data are publicly available at http://statewidedatabase.org/.

[12]Proposition 7 would have required California utilities to produce half their electricity from renewable

0.1% of blocks with the greatest percentage of individuals voting along green-friendly lines and a score near zero to the least green-friendly census block. Averaging these ordinal scores together yields the GII and maintains the 0 to 100 scale across all of California (the greenest areas are consistently green across all votes). We match individual households to the GII via their census block as a proxy for their own ideology under the assumption that like-minded individuals tend to cluster around green-related amenities that cater to their preferences.

2.4.3 Utilities

The three pilots provide an opportunity to estimate the effect of social-norm messaging in various climates and within communities with different demographic characteristics. For confidentiality concerns, we will refer to the utilities by the pseudonyms: Watertown, Aquaville, and Hydroburg. footnoteWaterSmart partners with water utilities who prefer that their names are kept anonymous when presenting data on consumer water consumption. Table 2.1 illustrates basic differences between the three pilots. Watertown and Aquaville are located in more moderate climates than Hydroburg; therefore water demand exhibits a different seasonal component. Since seasonality in water demand is primarily manifested through outdoor use, such as maintaining landscapes and refilling pools, the mechanisms by which water conservation programs affect consumption may vary by climate. The pilots also vary by average income and environmental ideology - both of which have an effect on water consumption. Hydroburg, the warmest, richest, and least environmentally conscious pilot has the highest baseline consumption. Watertown, the wettest, poorest, and greenest pilot, consumes the least water on average. A priori, these differences across pilots lead us to expect varied responses to WaterSmart's social-norm messaging program. For example, Costa and Kahn (2013) find that the response to social norms in energy is different for liberals and

resources by 2025. Proposition 10 would have allocated $5 billion as cash incentives for high fuel economy and alternative fuel vehicles and R&D for and education on renewable energy and alternative fuel technologies. Proposition 21 would have increased vehicle license fees in the state by $18 in order to raise $500 million a year dedicated to California State Parks. Proposition 23 would have suspended the Global Warming Solutions Act of 2006 until the state's unemployment rate decreased to 5.5% for four consecutive quarters, a condition not met for decades.

conservatives. In the case of income, more affluent consumers are more able to make larger water saving capital investments, however poorer households have a stronger financial incentive to reduce consumption because their water bill is a higher proportion of their disposable income.

Table 2.1: Summary Statistics By Pilot

Pilot	Water Use	Temp (F)	Rain (in)	Income	Home Value	Green Index
Watertown	202	70	2.2	57,333	358,394	71
Aquaville	282	67	1.7	96,696	670,242	62
Hydroburg	348	72	0.8	120,943	815,231	34

Notes: The water use reported is baseline water consumption in gallons per day; Temperature is the average monthly maximum temperature in degrees Fahrenheit; Precipitation is the monthly average in inches, median income is averaged at the census block group, and the green index is averaged at the census block.

The utilities also differ drastically in the characteristics of their service area. Aquaville has the largest customer base, serving over 1 million people. Due to its size there is greater variation in both climatic zones and demographic characteristics. Hydroburg is also a relatively large utility with over 300,000 customers, while Watertown serves less than 10,000 people. In addition to differences in the customer base, as evidenced in Table 2.1, utilities of such different sizes do not have the same resources at their disposal. Luckily, they do have similar existing conservation programs. Table 2.2 shows the various conservation programs in the three pilots *in addition* to WaterSmart's program. All the utilities offer rebate programs, as well as other programs tailored to the community's needs.[13]

The three utilities have very different rate structures that may influence treatment response in different ways. Watertown has the simplest rate with a fixed cost and single volumetric charge for the duration of the sample.[14] Aquaville has a standard increasing block rate structure with three tiers. The rate structure is relatively flat, with a 51% in-

[13]Hydroburg offers programs to address outdoor water use, while Watertown focuses on community engagement.

[14]During the treatment period plans were in place to move to an increasing block rate tariff.

crease from the lowest to the highest tier. Hydroburg has the most complicated and steepest rate structure. The rate is determined by the percentage of water a household uses relative to an allocation - determined by irrigable area and occupancy. The rate for the highest tier, reserved for consumption that is more than 200% above allocation, is over 600% higher than the base rate providing a strong financial incentive for conservation. Since the allocation acts as a signal for the appropriate level of use Hydroburg's water rate is structured to penalize households who use relatively more water than similar households. Therefore social norm messaging that highlights relative consumption resembles Hydroburg's rate structure, and may have a different impact compared to pilots where the rate structure does not reflect relative use. Since water demand is a function of utility-level policies, such as water rates and conservation programs, as well as the demographics and preferences of the consumer pool, we focus on the identification of treatment heterogeneity within pilots. Differences across pilots will be left to qualitative assessments driven by the empirical results. We leave the analysis of interaction between demand elasticity and social norms for future research.

Table 2.2: Utility Conservation Programs

	Watertown	**Aquaville**	**Hydroburg**
Rebates			
Toilets	X	X	X
Clothes Washer	X	X	X
Lawn Conversion	X	X	X
Sprinklers			X
Irrigation Controller		X	X
Technical Advice			
Home Water Audits	X	X	
Community Classes	X		

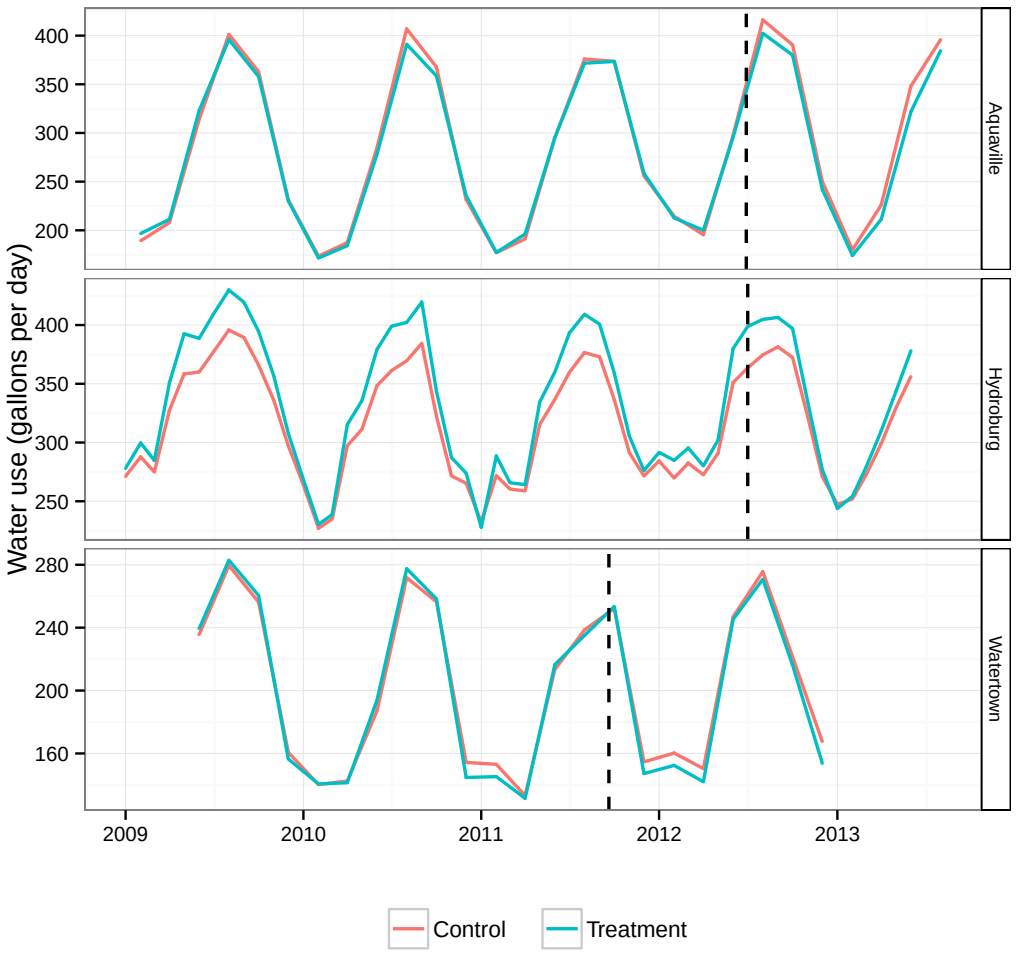
Notes: Only programs that are available from the utility are included. Some programs are administered through regional bodies, and additional resources are available from state and regional agencies.

2.4.4 Experimental Design

This section serves to describe the experimental design, defend the randomization process, and motivate the ability to identify causal effects and estimate several measures of treatment heterogeneity. Randomization eliminates concerns over selection bias, but it is important to evaluate the differences between the treatment and control group to determine the effectiveness of the randomization. Sample sizes, start dates, and the number of households treated in these pilots can be found in Table 2.4. We did not have control over the randomization process and the sample sizes are small enough to warrant inspecting the balance in observable variables across the treatment and control groups. Table B.16 displays the difference in means and the associated p-values from t-tests for several important variables. There are some variables that do have significant differences that warrant further examination, the most important being baseline water use in Hydroburg. Ensuring balance over the observable variables that are expected to influence heterogeneity in the treatment effect is an important prerequisite for an investigation into heterogeneous treatment effects. To ensure balance we employ a matching algorithm, described in the Data Appendix. Household fixed effects in our preferred specification will absorb all of the difference in household-level variables, but it is still important to assess if the treatment and control group have the same trend over time. Figure B.8 shows the time series of water use in gallons per day for both the treatment and control groups. It is clear from the graphs that all the pilots have similar trends over time across treatment status. This alleviates concerns about the significant difference in mean baseline water use in the estimation of the average treatment effect for Hydroburg even without matching. (See the Appendix for water use time series for each pilot in terms of year-over-year percentage changes.)

One avenue for addressing heterogeneity that is gaining attention in both applied, and theoretical, literature is the estimation of quantile treatment effects (QTE) (Fipro, 2007; Firpo et al., 2009; Allcott, 2011b; Ferraro and Miranda, 2013; Borah and Basu, 2013; Frölich and Melly, 2012, 2013). This is important for policymakers who want to know the type of

Figure 2.3: Baseline Water Use by Pilot



Notes: The graph contains the mean water used per day in gallons for both the treatment and control groups over time. The vertical dashed line indicates the start of the program for each pilot.

Table 2.3: Differences in Means Between Control and Treatment Groups

(a) Variables used to explore heterogeneity

Pilot	Baseline Water	Green Index	Home Value	Income
Aquaville	-1.211 (0.861)	0.3 (0.48)	-12,107 (0.946)	-84 (0.77)
Hydroburg	13.89 (0)	2.7 (0.606)	109,701 (0.979)	3,337 (0.011)
Watertown	-6.541 (0.993)	0.5 (0.195)	-19,930 (0)	-530 (0.21)

(b) Structural Features of the House

Pilot	Sqft	Lot Size	Single Family Home	Year Built
Aquaville	-11 (0.753)	-178.4 (0.116)	0 (0.112)	1.7 (0.538)
Hydroburg	208 (0.292)	217 (0.738)	-0.015 (0)	8.5 (0.495)
Watertown	-32 (0.123)	-124.6 (0.854)	-0.029 (0.126)	1.2 (0.079)

Notes: Values reported are the control mean minus the treatment mean. P-values for the difference from zero are in parenthesis. Sqft is the square feet of the home, the Lot Size is measured in square feet, Single Family Home is a proportion of the sample in a detached house, and Year Built is the year the home was built. While we don't use income explicitly to explore heterogeneity, the use of home value is used as a proxy because of its finer geographical granularity.

Table 2.4: Sample Sizes

Pilot	Start Date	End Date	N: Obs	N: Post-Treat	HHs	Treated HHs
Watertown	2011-09-20	2013-01-01	44,822	16,266	2,575	992
Aquaville	2012-06-28	on going	93,664	18,583	3,096	1,547
Hydroburg	2012-07-01	on going	666,981	115,649	11,710	1,184

households that are most responsive to an intervention and whether a program may have the perverse effect on certain segments of the population. The aforementioned literature highlights some of advantages of unconditional QTEs that provide inference on the marginal distribution of the outcome variable, as opposed to the conditional distribution. The conditional quantile does not have the same interpretation as the conditional mean, and in general the marginal distribution is more relevant in a policy context. The interpretation of the QTE is subject to the rank preservation assumption that presumes treatment assignment does not affect the relative ranking of households consumption. If the assumption holds the process

of treatment assignment does not systematically determine the consumption quantile for a household. Assuming rank preservation holds a control household at a given quantile is an appropriate counterfactual for the corresponding treated household. Therefore the QTE represents quantiles of the treatment effect, and provides inference on the distribution of treatment effects across households at different quantiles. Without this assumption, we can only interpret the results from quantile regressions as the QTE on the outcome distribution (Fipro, 2007), which still may have economic and policy relevance.

It is not possible to directly test this assumption, but the literature (Bitler et al., 2006, 2008; Djebbari and Smith, 2008) suggests an ad hoc assessment that examines the difference in means within quartiles across treatment status for key variables.[15] Many significant differences in means make the rank preservation assumption difficult to justify since the distribution of observables that affect water demand vary across treatment status. However, even good balance within quartiles does not prove that the rank preservation assumption holds. Table 2.5 reports the p-values for difference in means for several variables within quartiles for each pilot. 25% of the differences have p-values less than 0.05, suggesting that the quantile treatment effects may provide a decent approximation for the quantiles of the treatment effect. In the next section we present our estimation strategy and empirical results.

[15] All the variables in this test are pre-determined relative to treatment assignment, and this primarily motivates that households at a given quantile have similar characteristics.

Table 2.5: Motivating the Rank Preservation Assumption

Pilot	Quantile	Baseline Water	Green Index	Lot Size	Sq.Ft.	Year Built
Aquaville	1.0	(0.352)	(0.102)	(0.013)	(0.111)	(0.146)
Aquaville	2.0	(0.003)	(0.006)	(0)	(0.747)	(0.138)
Aquaville	3.0	(0)	(0.624)	(0.001)	(0.939)	(0.131)
Aquaville	4.0	(0.742)	(0.861)	(0.109)	(0.552)	(0.959)
Hydroburg	1.0	(0)	(0.303)	(0.195)	(0.001)	(0.368)
Hydroburg	2.0	(0)	(0.187)	(0.895)	(0)	(0.04)
Hydroburg	3.0	(0.002)	(0.001)	(0.318)	(0.086)	(0.812)
Hydroburg	4.0	(0.262)	(0.942)	(0.581)	(0.489)	(0)
Watertown	1.0	(0)	(0)	(0.005)	(0.127)	(0.243)
Watertown	2.0	(0)	(0)	(0)	(0.029)	(0.273)
Watertown	3.0	(0)	(0)	(0)	(0)	(0.82)
Watertown	4.0	(0.782)	(0.049)	(0.415)	(0.572)	(0.149)

Notes: Table contains the p-values for difference in means between treatment and control groups within each quartile. The rows designate a pilot-quartile pair and the the columns are variables important to water consumption.

2.5 Estimation Strategy

2.5.1 Average Treatment Effects

The primary goal of this research is to estimate the average treatment effect and explore treatment effect heterogeneity. Because households were randomly selected into treatment, a simple t-test for difference in water use means after treatment between the treatment group and the control group provides a valid statistical tool for the analysis. However, water demand has a large variance and in order to identify relatively small effect sizes with the conventional significance levels we need to control for covariates.[16] A Hausman test that is robust to heteroskedastic and clustered standard errors as suggested by Cameron and Trivedi (2005) provides evidence for the fixed effects model written as

$$\ln(w_{it}) = \alpha_i + \gamma T_i \times P_t + \theta P_t + \beta X'_{it} + \xi_{it} \tag{2.1}$$

[16]The conventional for a small effect size is 0.2 (standardized to the variance of the outcome variable) and in our setting we are attempting to identify effect sizes on the order of 0.05, which is only possible while controlling for covariates.

where w_{it} is water consumption for household i at time t, T_i is a dummy that identifies the treatment group, P_t is a dummy for treatment period, X_{it} is a set of time-varying covariates, and ξ_{it} is an idiosyncratic error term. In our regressions the covariates include the average number of rainy days and cooling degree-days; as well as year-period fixed effects. Huber-White standard errors (Woolridge, 2002) are clustered at the household level to account for serial correlation within a household over time. The regression results for each utility are presented in Table 2.6.

Table 2.6: Average Treatment Effects

	(1) Watertown	(2) Aquaville	(3) Hydroburg
Treatment Effect	-0.0534***	-0.0313***	-0.0170**
	(0.0146)	(0.0119)	(0.0073)
Rainy Days	24.1253	-0.3020***	-8.5027***
	(19.8758)	(0.0367)	(0.3942)
Cooling Degree Days	0.2633	0.0067***	0.0013
	(0.6321)	(0.0009)	(0.0047)
Constant	4.6138***	5.1315***	5.5719***
	(0.0972)	(0.0130)	(0.0046)
Household FEs	Yes	Yes	Yes
Year-Period FEs	Yes	Yes	Yes
Adjusted R^2	0.157	0.251	0.190
Households	2,575	3,093	11,696
Observations	44,763	83,342	664,075

Notes: The dependent variable is the natural logarithm of water consumption in gallons per day. Rainy days and cooling degree days are the average during the reading period, generated from daily weather data. Robust standard errors clustered at the household level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

We find that sending home water reports reduces average monthly water use in the billing periods after treatment by 5.34% in Watertown, 3.13% in Aquaville, and 1.70% in

Hydroburg.[17] The results from Table 2.6 display the differences in treatment response across pilots, and we run a pooled regression with pilot-level interactions for all variables to test if these differences are statistically significant. An F-test rejects the null for a constant ATE across pilots at the 1% level. In grouped F-tests between each pair of pilots we reject the null for equal ATEs for Watertown and Hydroburg at the 5% level and for Hydroburg and Aquaville at the 10% level. In determining whether to run a pooled model with pilot-level interaction terms we must evaluate whether it is appropriate to pool the error term across all the pilots. One way to evaluate if unobserved heterogeneity is equal across pilots is to examine the observed heterogeneity. All the interaction terms for weather variables and many of the interactions for year-period fixed effects are statistically significant. The presence of significant observed heterogeneity signals a likelihood that unobserved heterogeneity such as utility policies and regional macro-level shocks exists that will cause the error variances to differ across pilots. The results suggest that it will be difficult to identify pilot-level effects in a pooled model, and therefore we restrict our econometric analysis to within-pilot heterogeneity.

2.5.2 Durability

Our first foray into the heterogeneity looks at the treatment effect over time. Ex ante there are several valid hypotheses explaining the dynamic pattern of consumer response to the treatment. One potential pattern of results is that initially the program increases the salience of water consumption and households make temporary behavioral adjustment to their use. In this setting, after an initial strong response, the treatment effect will wane over the course of the program. Another conceivable outcome is that the program induces capital investments or permanent changes in behaviors that cause consistent reductions throughout the program. The series of period-level treatment effects over time also provides insight into the mechanisms through which the HWRs cause consumers to reduce water use, including

[17] Relative effect of a dummy variable with estimated coefficient α in a log-linear regression is equal to $e^{\alpha} - 1$ (Halvorsen and Palmquist, 1980).

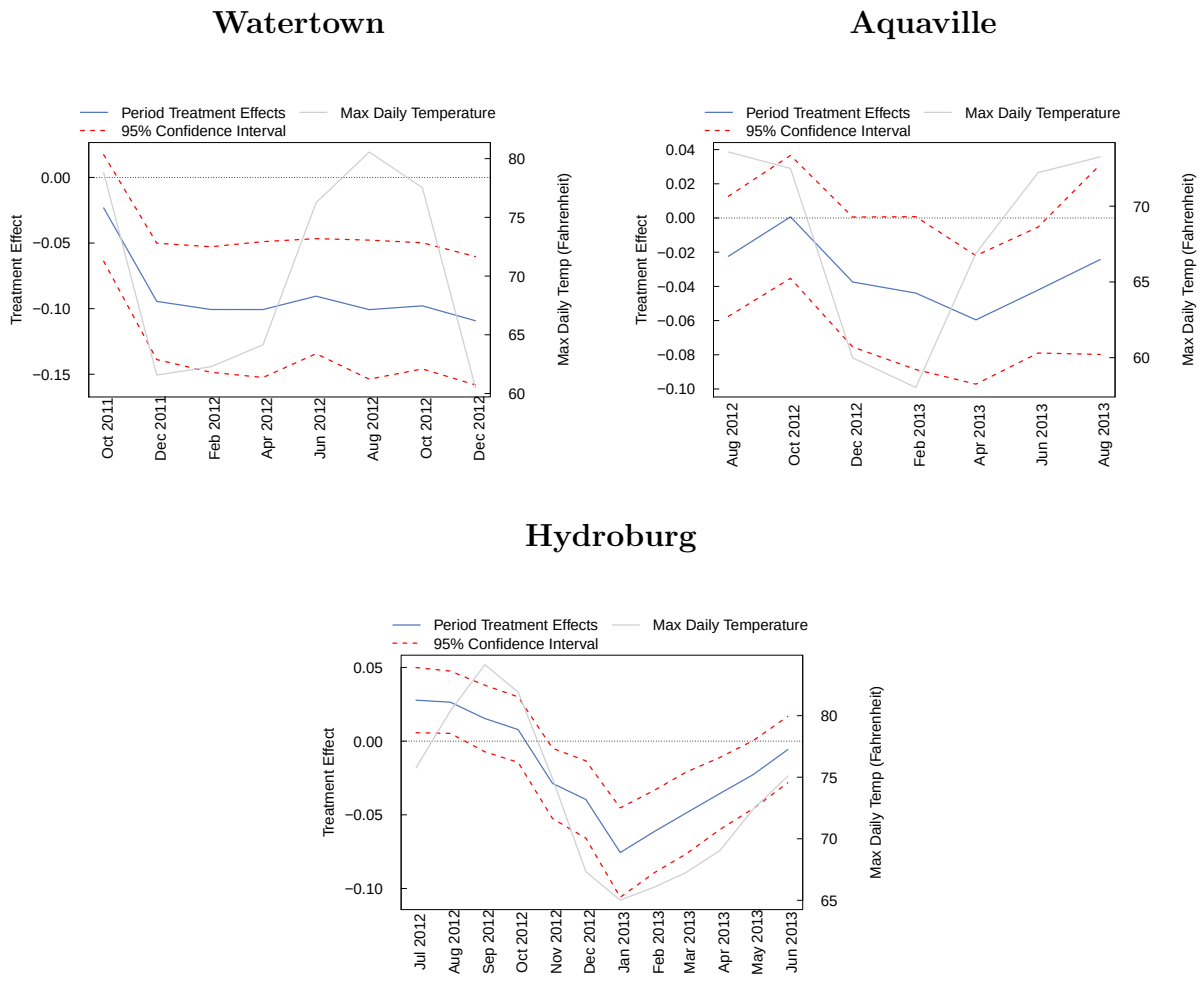
the role of seasonality and outdoor water demand. Both Ferraro and Price (2013) and Allcott and Rogers (2012a) examine the effect of social norm messaging after the treatment ceases. Allcott and Rogers (2012a) distinguish "durability", the effect over time during the treatment period from "persistence", how the effect remains after treatment ends. They find that, in energy, the effect of social norms is both durable and persistent, though the persistence wanes with time from the end of treatment. In Ferraro and Price (2013) the social norm is only sent in one summer so analysis of durability is not possible. They see a similar tapering in the persistence of the treatment effect over time as Allcott and Rogers (2012a). Our experimental design allows us to examine durability but not persistence, because the control group also begins receiving the treatment (the HWR) at the end of the experiment. In order to examine the durability of the treatment effect *within* the active administration of the program we run separate regressions for each treatment period. Thus the regression equation is now

$$\ln(w_{it}) = \alpha_i + \gamma T_i \times P_{t,h} + \theta P_{t,h} + \beta X'_{it} + \xi_{it} \tag{2.2}$$

where $h = 1, ..., H$ is a specific treatment period as opposed to the whole course of the program. In this setting we drop all observations for other treatment periods, P_{-h}, so that we compare a single treatment period to all controls as opposed to other treatment periods. The results from equation 2.2 for all treatment periods and for each pilot are shown graphically in Figure 2.4. The left vertical axis represents the coefficient and 95% confidence intervals and the horizontal axis is the series of treatment periods over time. To highlight the interaction with seasonality we include average pilot-level evapotranspiration, which measures the consumptive use of turf grass in millimeters, on the right vertical axis.

The graphs reveal variation between pilots in the pattern of savings over the course of the program. Note that Watertown and Aquaville have meters read every two months so each period corresponds to two calendar months, whereas in Hydroburg the time series is observed every month. The first home water report for each pilot is sent out in the middle of the reading period, after the previous period's data is processed in order to generate the

Figure 2.4: Durability of Treatment Effects - Percentage



Notes: The solid line is set of point estimates for a regressions on an individual treatment period. All data from other treatment periods are omitted and all regressions contain household and year-period fixed effects as well as weather controls. The dashed line represents the 95% confidence interval constructed from robust standard errors clustered at the household level. The grey line, on the right vertical axis, is the average maximum daily temperature in degrees Celsius for the pilot to highlight the correlation of individual-period treatment effects and weather.

metric on relative consumption. Therefore, a portion of consumption in the first period is actually untreated, which attenuates the treatment effect. This explains the smaller savings in the first period relative to most other periods observed in all pilots. As a robustness check we re-run the regressions from equation 2.1 dropping all observations in the first treatment period for each pilot. These ATEs, as seen in Table 2.7, are very similar to the base regressions in Table 2.6 but are slightly larger in absolute value(-5.46%, -3.31%, and -2.02% for Watertown, Aquaville, and Hydroburg, respectively).

Table 2.7: Average Treatment Effects: Exclude first period

	(1) Watertown	(2) Aquaville	(3) Hydroburg
Treatment Effect	-0.0561***	-0.0337***	-0.0204***
	(0.0148)	(0.0118)	(0.0076)
Rainy Days	23.8145	-0.3048***	-7.6478***
	(19.8971)	(0.0370)	(0.3997)
Cooling Degree Days	0.2959	0.0069***	0.0121**
	(0.6425)	(0.0009)	(0.0047)
Constant	4.8397***	5.1302***	5.5307***
	(0.0804)	(0.0129)	(0.0047)
Household FEs	Yes	Yes	Yes
Year-Period FEs	Yes	Yes	Yes
Adjusted R^2	0.164	0.253	0.183
Households	2,504	3,092	10,490
Observations	42,394	80,845	589,030

Notes: The dependent variable is the natural logarithm of water consumption in gallons per day. All observations from the first treatment period for each pilot are omitted. Rainy days and cooling degree days are the average during the reading period, generated from daily weather data. Robust standard errors clustered at the household level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The treatment effect for both Watertown and Aquaville is relatively stable after the first period (Figure 2.4), suggesting that consumers either made early investments in water efficiency or established permanent behavioral changes. The treatment effect in Hydroburg

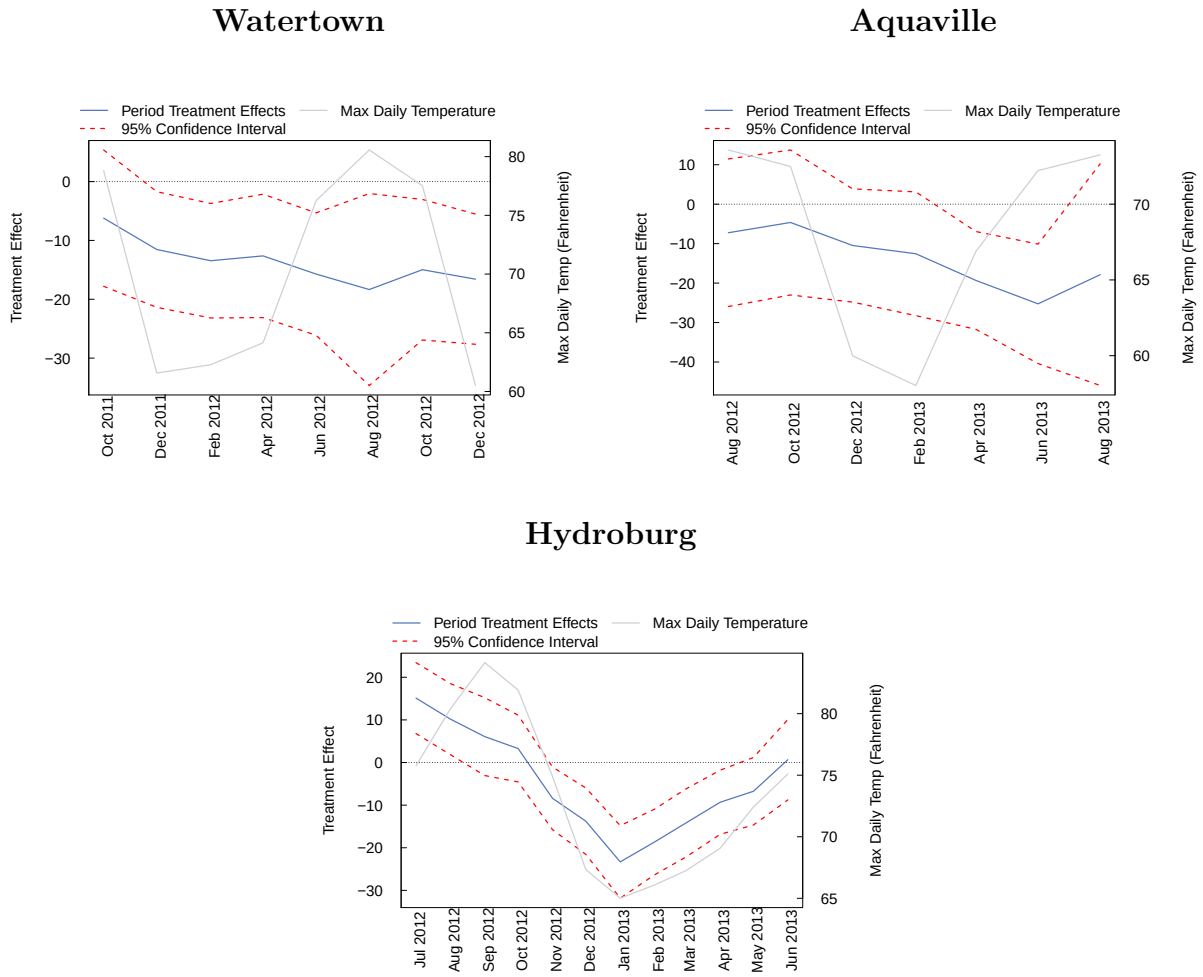
took several months to appear (perhaps because there was a lag between initially receiving the HWR and undertaking conservation efforts), and then wanes over the treatment period. The treatment effect is essentially zero in Hydroburg 11 or 12 months after the HWRs were first sent. This lack of durability in the treatment in this pilot should not be confused with persistence; the HWRs continued to be sent to households in Hydroburg through June 2013.

Learning whether the program is effective at reducing outdoor water use is critical. This is the peak demand period and often corresponds with supply shocks. If the main saving mechanism was in reducing outdoor water use we would expect a negative correlation between temperature and the treatment effect because larger reductions in water use would correspond with higher temperatures. Examining the pattern of individual treatment effects in conjunction with the weather data indicates that the social norm does not generate savings primarily through outdoor use (Figure 2.4). This result is not surprising given that Watertown has the most moderate climate and relatively less outdoor use than the other pilots. However, even in Hydroburg, the hottest and driest pilot, the correlation is positive with peak savings in January- exactly the opposite result that would support savings in outdoor water use. Similarly, the U-shaped curve for Aquaville suggests that more conservation occurs in the winter months with peak savings occurring in March and April.

Though we control for seasonality with year-period fixed effects and weather variables, there may be concern that the logarithmic transformation masks some of the aggregate savings in gallons. Since water use is higher in the summer the same amount of gallons saved during the winter is larger relative to average in percentage terms. In order to test the sensitivity to our specification of the dependent variable we re-run the regressions in equation 2.2 with water use in gallons per day. Investments made in indoor water efficiency during the winter should still be visible in later summer months. The results are graphed in Figure 2.5. In Aquaville we observe a steadily increasing treatment effect in absolute value (except for a slight decrease in July 2013), supporting the notion that households gradually invest in water efficiency after receiving the HWR. However, regressions on water as opposed to log water

do not change the pattern in Hydroburg, suggesting a different experience. The return to a zero treatment effect after several months suggests transient behavioral savings as opposed to investments in efficient consumer durables. Hydroburg's existing rate structure already reflects relative water use and is very punitive for high relative water users. Consumers in this pilot may have already made investments in water efficiency that were not yet undertaken in the pilots with less aggressive rate structures.

Figure 2.5: Durability of Treatment Effects - Gallons per day



Notes: The solid line is set of point estimates for a regressions on an individual treatment period. All data from other treatment periods are omitted and all regressions contain household and year-period fixed effects as well as weather controls. The dashed line represents the 95% confidence interval constructed from robust standard errors clustered at the household level. The grey line, on the right vertical axis, is the average maximum daily temperature in degrees Celsius for the pilot to highlight the correlation of individual-period treatment effects and weather.

2.5.3 Treatment Effect Heterogeneity

There are many ways to explore the treatment effect heterogeneity for interventions using social norms to promote sustainable behavior. Allcott (2011b) use quantile regression and interactions with baseline electricity consumption quantile dummies, Ferraro and Miranda (2013) explore the baseline consumption and differences due to home prices, and Costa and Kahn (2013) utilize voting data to examine differences in treatment effect due to ideology. We build on the prior research by combining these approaches in a setting where we observe considerable variation in both ideology and treatment response.

Quantile Treatment Effects

We estimate unconditional quantile regressions in the presence of covariates through the methodology of Firpo et al. (2009). Estimating conditional quantile regressions in the vein of Koenker and Basset (1978) is less appealing in a policy framework when the unconditional distribution is of primary interest. Interpreting a coefficient from a quantile regression when conditioning on many variables is difficult and the results can vary substantially depending on the data generating process (Borah and Basu, 2013). Since our data measure water use over time, the identification of the treatment effect requires controls for time varying covariates through weather and time fixed effects.[18] The widely employed estimator from Fipro (2007) only introduces covariates to remove selection bias in estimating quantile treatment effects, a problem that does not exist in our case of random assignment into treatment. Therefore we use the estimator of Firpo et al. (2009) in order to identify treatment effects by controlling for time varying covariates while still providing unconditional quantile treatment effects, the quantiles of the treatment effect over the distribution of all households under current assumptions.

The basic logic behind the estimator draws from the influence function (IF) describing

[18]If our samples sizes were larger we should be able to identify the treatment effects with a simple difference in means, or difference in quantiles, but we require covariates to identify statistically significant effects.

the influence of an individual observation on a given distributional statistic. A convenient feature is that the recentered influence function (RIF) is simply the IF plus the distributional statistic; the expectation of the RIF is the distributional statistic itself. In the case of quantiles the RIF for the τth quantile is $q_\tau + (\tau - \mathbb{1}\{Y \leq \tau\}/f_Y(q_\tau)$. The RIF regression function with covariates is $E[RIF, Y, q_\tau|X] = m_\tau(X)$, which Firpo et al. (2009) prove represents an unconditional quantile regression. Given the assumption on the validity of the control group as a counterfactual the coefficient representing the treatment effect can be interpreted as the QTE. With the rank preservation assumption, the coefficient can be interpreted as the quantile of the treatment effect.

Even without the rank preservation assumption, learning about the difference in marginal distributions of water consumption between treated and untreated groups may be relevant to policymakers. One important clarification is that the estimator does not account for the panel structure of the data and contains no household fixed effects. Exploring quantile treatment effect models that incorporate the panel data structure and household fixed effects is left for future research and will follow Powell (2013).

We estimate the QTE using the RIF approach at each quantile of water use from 0.05 to 0.95 in 0.05 intervals with standard errors that are robust to heteroskedasticity. Figure 2.6 shows the results for the quantile treatment effects. The thick blue line represents the coefficient estimate for the treatment effect at each quantile of water use. The thick red dashed lines are the 95% confidence interval. The corresponding OLS estimates and 95% confidence interval are also shown on the graph by the horizontal green dashed and dotted lines respectively. The QTE is relatively constant across quantiles for Hydroburg. Aquaville and Watertown show a similar pattern to Allcott (2011b) and Ferraro and Miranda (2013) in that the QTE increases at the highest quantiles. Watertown is unique in that there is a significantly non-zero QTE at low quantiles as well. This is a different pattern than seen in the literature on the effect of social norms and argues for the effectiveness of social norms at the low end of the consumption distribution. Another key finding is that at no point

in the distribution is the QTE above zero, which is important for policymakers cautious of provoking the perverse effect of increased consumption in response to messages of relatively low water use.

In Figure 2.6 Watertown does experience savings at the lower end of the consumption distribution with a significant QTE at $tau = 0.1$, however the results suggest that targeting large water users will generate larger savings per household. Based on the empirical results WaterSmart is concentrating an expansion in one pilot exclusively for high water users.[19]

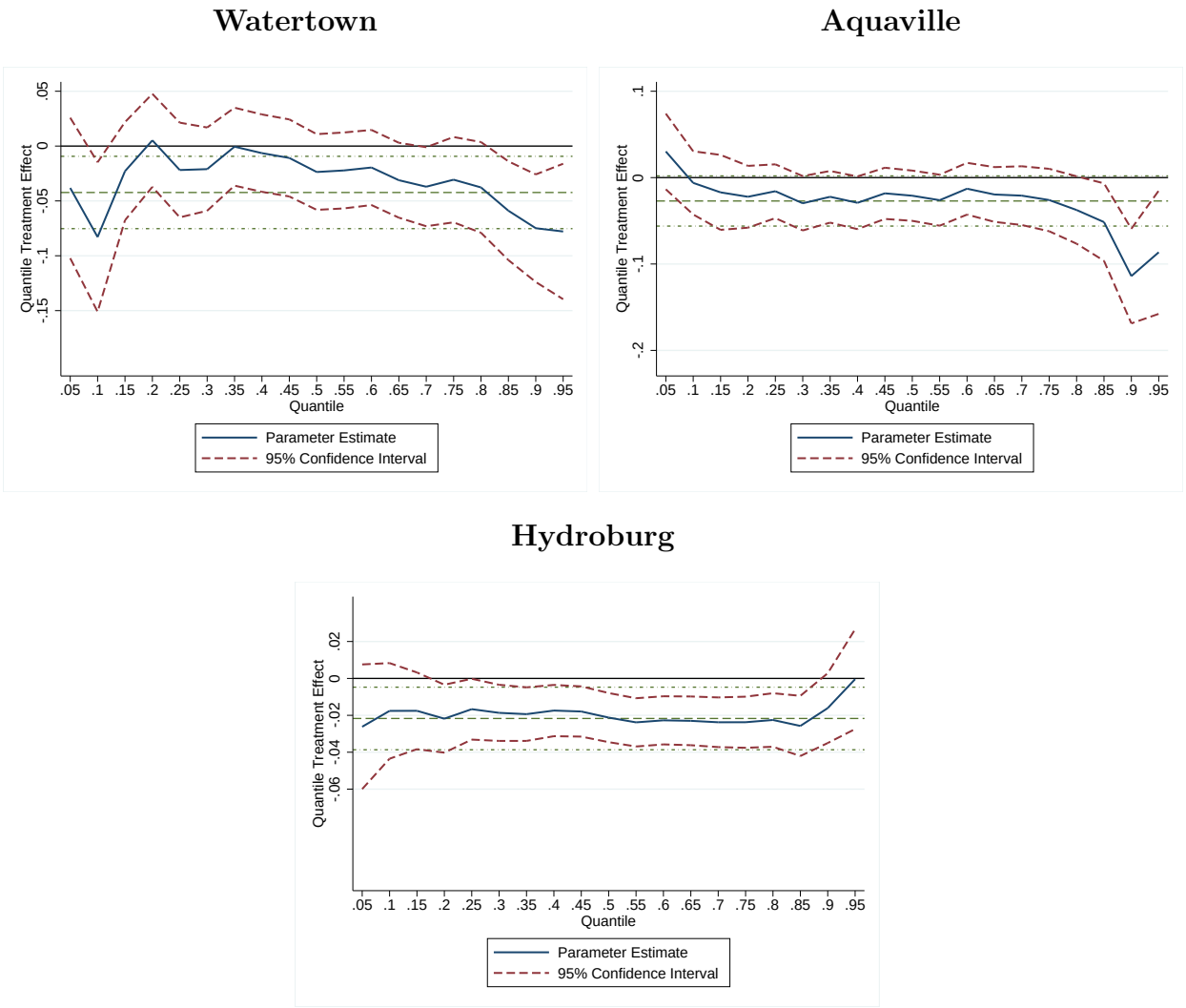
Conditional Average Treatment Effects

For a more intuitive regression model in addition to the unconditional quantile regressions above, we interact the treatment effect with quintiles of baseline water consumption for each utility. These regressions also attempt to parse out the effect of treatment across the distribution of water use, but focus on conditional average treatment effects (CATE) for each quintile as opposed to QTE. In this case the treatment effect of higher quantiles are only relevant for household types that populate that end of the distribution. Baseline deciles are calculated by taking the average pre-treatment household usage across the whole utility and ordering it into equal groups of five. We perform a similar technique to create quintiles of the GII and housing values for each utility. Note that higher quintiles of baseline consumption relate to higher baseline usage and higher GII quintiles correspond to stronger environmental preferences. The regressions include all the variables as presented in equation 2.1 in addition to interactions between the treatment effect and the quintiles of water and GII. The regressions have the form

$$\ln(w_{it}) = \alpha_i + \gamma T_i \times P_t + \sum_{d=1}^{5} \delta_d (T_i \times P_t \times D_d) + \theta P_t + \beta X'_{it} + \xi_{it} \tag{2.3}$$

[19]The targeted expansion is occurring in the population that does not have email addresses registered with the utility company, requiring a more costly printed version of the HWR to be sent by conventional mail as opposed to email.

Figure 2.6: Quantile Treatment Effects



Notes: The thick blue line is the estimate for unconditional quantile treatment effect by quantile in 0.05 intervals, with red thick dashed line representing 95% confidence intervals based on heteroskedastic-robust standard errors. The OLS regression and 95% confidence intervals are the flat horizontal dashed and dotted lines respectively. Pooled OLS is used for comparison, as opposed to the FE model, because the quantile regressions do not account for the panel structure of the data.

where D_d is the d^{th} quintile of either baseline consumption, the GII, or housing values. The parameters of interest are the linear combination of γ and δ_d.

Table 2.8 shows the effect of treatment for households in different quintiles of baseline water use. The estimates in Table 2.8 are for $\gamma + \delta_d$, the linear combination of the treatment effect and the interaction of the treatment effect with each quintile of baseline consumption.[20] The significance levels are determined by p-values from test for joint significance using F-statistics. The full regression output is provided in Table B.18 in the appendix. In Watertown the highest three quintiles are all negative and statistically significant, but the lower two quintiles are not significant. This is an indication that the majority of savings comes from households with above average water consumption. Similarly, Hydroburg has negative and significant coefficients on the top three quintiles, but also has a positive coefficient on the second quintile. This suggests that some households may actually increase consumption due to the treatment. For Aquaville the only statistically significant quintile is in the middle of the distribution. The lowest quintile in Aquaville is positive suggesting a similar relationship seen in the other pilots that the program is less effective in households that start out using less water. This raises the question of whether the social norms are effective for households that already are consuming below the average prior to the intervention. This results are slightly different from those conclusions drawn from the unconditional quantile regressions, with starker differences between the two methodological results for Aquaville and Hydroburg.

Unlike the impact of baseline consumption, the effect of ideology and housing values on treatment effect heterogeneity is not consistent across the three pilots. Tables 2.9 and 2.10 show the linear combinations of of treatment effect and interaction terms for regressions based on quintiles of the GII and housing values. Watertown displays relatively consistent CATE across all quintiles of the GII and housing values, though there are insignificant effects at the second quintile for both variables. Aquaville and Hydroburg show opposite effects for both variables, though the interactions are generally insignificant and it is difficult to discern

[20]The third quintile, containing the median, is the omitted category and in this case the raw coefficient and stand error are shown.

Table 2.8: Heterogeneity: Baseline Water Use

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Watertown	-0.00502	-0.0104	-0.0537*	-0.0853***	-0.110***
	(0.0289)	(0.0233)	(0.0237)	(0.0211)	(0.0195)
Aquaville	0.0213	-0.0345	-0.0732***	-0.0361	-0.0366
	(0.0225)	(0.0219)	(0.0195)	(0.0191)	(0.0210)
Hydroburg	0.0208	0.0439**	-0.0509***	-0.0449***	-0.0415**
	(0.0192)	(0.0158)	(0.0138)	(0.0132)	(0.0149)

Notes: The coefficients reported here are the linear combinations of the coefficients from the regression results in Table B.18. Robust standard errors clustered at the household level for these linear combinations are reported in parentheses. Stars denote significance from F-tests evaluating the difference of the linear combination from zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

a trend. For ideology there is a positive, but insignificant effect at the second quintile in Aquaville and the fifth quintile in Hydroburg. Interpreting the results with caution due to the noise in the estimate, it appears that more environmentally-conscious households respond more to social norms in Aquaville while in Hydroburg the least eco-friendly households are more responsive. One caveat is that the ideology data are at the census block level as opposed to the household level as in Costa and Kahn (2013).

In terms of housing values, which serve as a proxy for income, Aquaville has negative and significant effects at the lowest two quintiles whereas Hydroburg has a significant effect at the fourth quintile. The results from CATEs by housing value Aquaville are the starkest, and also contradict the finding of Ferraro and Miranda (2013) that households in expensive homes are more responsive. Overall it is difficult to make any conclusions about how environmental ideology and housing values drive heterogeneity in the response to treatment. Most of the interaction terms as seen in Tables B.19 and B.20 in the Appendix are insignificant, and the trends run in opposite directions for Aquaville and Hydroburg. This suggests that the biggest driver of heterogeneity is baseline water consumption. Another issue is that we are only exploiting intra-pilot variation in ideology and housing values. There is more

variation across pilots for these two variables, and that may be part what is determining the difference in magnitudes of the pilot-level ATEs. In a purely descriptive analysis however, the ATE of social norm messaging in Hydroburg the most conservative and affluent area in our study, is lower than that of the other two pilots. Watertown, the least affluent and most environmentally friendly utility had the strongest ATE. We must note here that Hydroburg's long standing rate structures aggressively tied to relative water use confound even these descriptive conclusions.

Table 2.9: Heterogeneity: Ideology

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Watertown	-0.0682**	-0.0271	-0.0533	-0.0634*	-0.0930**
	(0.0252)	(0.0262)	(0.0281)	(0.0317)	(0.0311)
Aquaville	-0.0152	0.0196	-0.0458*	-0.0258	-0.0516
	(0.0211)	(0.0215)	(0.0231)	(0.0225)	(0.0291)
Hydroburg	-0.0234	-0.0679***	-0.0148	-0.0184	0.00879
	(0.0167)	(0.0163)	(0.0156)	(0.0171)	(0.0210)

Notes: The coefficients reported here are the linear combinations of the coefficients from the regression results in Table B.18. Robust standard errors clustered at the household level for these linear combinations are reported in parentheses. Stars denote significance from F-tests evaluating the difference of the linear combination from zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.10: Heterogeneity: Housing Values

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Watertown	-0.0662*	-0.0171	-0.0593**	-0.0621**	-0.0603**
	(0.0272)	(0.0268)	(0.0216)	(0.0232)	(0.0219)
Aquaville	-0.0704**	-0.0701**	-0.0372	-0.0108	0.0196
	(0.0232)	(0.0252)	(0.0216)	(0.0185)	(0.0170)
Hydroburg	-0.0176	-0.0276	-0.0251	-0.0295*	-0.00622
	(0.0173)	(0.0180)	(0.0155)	(0.0138)	(0.0157)

Notes: The coefficients reported here are the linear combinations of the coefficients from the regression results in Table B.18. Robust standard errors clustered at the household level for these linear combinations are reported in parentheses. Stars denote significance from F-tests evaluating the difference of the linear combination from zero. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.5.4 Interactions with Existing Conservation Programs

Since the HWRs provide information about utilities' water conservation programs and generally increase the awareness of water consumption, we may expect those in the treatment group to participate in those programs at a higher rate. On the other hand, the deluge of information about water conservation may inundate consumers and social norms may crowd out existing conservation efforts. Using data on participation in conservation programs that spanned both the pre and post-intervention periods, we estimate the impact of the social norm message on the probability of engaging in an additional utility program.

In the three pilots only Watertown, the longest running pilot, had enough post-treatment participation data to estimate a statistical model. We hope to get sufficient data from the other pilots in the future.[21] We estimate a logit model where the dependent variable is a dummy equal to one if a household participates in a program in a given period and zero otherwise. We run several specifications of the dependent variable. The first specification is to pool all the different programs together to create a single indicator for participation, defined as all programs. We also create separate indicators for the completion of a home

[21] Aquaville and Hydroburg had many observations that were missing a program participation date. We will likely be able to run these models as the utilities enter recent participants into their database.

visit by an expert to provide personalized advice on water conservation, the receipt of a high-efficiency toilet rebate, or the receipt of a high-efficiency clothes washer rebate.[22] Since our goal is to assess the impact of the social norm campaign on participation and we only observe 384 instances of program participation during our sample, we simplify the temporal dimension to before or after treatment. Our regression then takes the form of a difference-in-difference logit model where we control for any time and group level effects as well as static household characteristics that may influence participation.

$$\Pr(c^l_{i,t} = 1|X_{i,t}) = \frac{\exp(\beta X_{i,t})}{1 + \exp(\beta X_{i,t})} \tag{2.4}$$

where $X_{it} = T_i + P_t + T_i \times P_t + Z_i$. The treatment group and treatment period indicator are defined as T_i and P_t respectively, i designates the household, time, $t = \{1, 2\}$, is pooled into pre and post intervention periods, l specifies the type of conservation programs and Z_i is a vector of household characteristics.

The results from equation 2.4 are presented in Table 2.11 as marginal effects, with standard errors calculated by the delta method. The treatment effect defined as $T_i \times P_t$ increases the probability of participation in conservation programs in the pooled specification as well as the completion of a home water audit, but is statistically insignificant for each of the individual rebates. These results provide evidence that social norms are a complement to existing programs and do not crowd out other conservation efforts.

In order to connect increased participation rates to water use we estimate the impact of the conservation programs on water demand. The regressions take the following form

$$\ln(w_{it}) = \alpha_i + \gamma T_i \times P_t + \pi_1 C^l_{i,t} + \pi_2 \tilde{C}^l_{i,t} + \theta P_t + \beta X'_{it} + \xi_{it} \tag{2.5}$$

where $C^l_{i,t}$ is the cumulative sum of programs that household i has participated in at time t, for l defined by the categories: all programs, home visit, toilet rebate, and clothes washer

[22]There are two programs, greywater systems and landscape conversion rebates, that are included in "all programs" category but not estimated separately due to the small number of observations.

Table 2.11: Logit Regressions for Conservation Programs - Watertown

	Pooled		Rebates	
	(1)	(2)	(3)	(4)
	All Programs	Home Survey	Toilet	Clothes Washer
Treatment Effect	0.1605***	0.1178***	0.0102	0.0001
	(0.0359)	(0.0363)	(0.0140)	(0.0085)
Treatment Group	0.0093	-0.0005	0.0019	0.0004
	(0.0080)	(0.0058)	(0.0016)	(0.0029)
Treatment Period	-0.1381***	-0.0308***	-0.0484***	-0.0339***
	(0.0161)	(0.0090)	(0.0146)	(0.0079)
Single Family Home	0.0102	0.0021	0.0032	0.0002
	(0.0094)	(0.0053)	(0.0020)	(0.0040)
Bathrooms	0.0245***	0.0065**	0.0042***	0.0045**
	(0.0048)	(0.0028)	(0.0016)	(0.0019)
Home Size (Sq. Ft.)	-0.0017	0.0023	-0.0010	0.0013
	(0.0098)	(0.0055)	(0.0020)	(0.0040)
Year Built	-0.0001	-0.0000	-0.0001**	-0.0000
	(0.0002)	(0.0001)	(0.0000)	(0.0001)
Lot Size (Sq. Ft.)	-0.0091**	-0.0112***	-0.0025**	-0.0007
	(0.0037)	(0.0029)	(0.0011)	(0.0008)
Home Value	0.0075	0.0024	0.0005	0.0039**
	(0.0048)	(0.0029)	(0.0011)	(0.0018)
Pseudo R2	0.0849	0.0859	0.156	0.106
Observations	4,015	4,015	4,015	4,015

Notes: The dependent variable is a dummy for participation in a given water utility conservation program. The data are pooled temporally by pre/post WaterSmart intervention. The marginal effects are reported along with standard errors obtained from the delta method.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

rebate. $\tilde{C}^l_{i,t}$ is the sum of programs for those in the treatment group that were initiated after the start of the social norm campaign. Since households can participate in multiple programs of the same type $C^l_{i,t}$ and $\tilde{C}^l_{i,t}$ not only capture the presence but also the intensity of

program participation. The term $\tilde{C}^l_{i,t}$, essentially an interaction term, represents the marginal effect of the normative message on water savings beyond the baseline impact of conservation programs and identifies whether information campaigns improve the efficiency of existing campaigns. Bennear et al. (2013) argues that conservation programs like rebates do not pass the test of additionality because those that take advantage of rebates may have undertaken the conservation activity regardless of the presence of the rebate. If a social norm campaign prompts households to engage in a toilet rebate who otherwise would not have replaced their toilet, then the conservation programs should be more effective in the population treated by the social norm. Due to random assignment into treatment we do not expect differences in natural appliance depreciation rates across groups. Therefore, if the social norm spurs additional program participation it is likely that these are households that otherwise would not have replaced inefficient appliances. In other words, a positive coefficient on $\tilde{C}^l_{i,t}$ means that households that receive a social norm utilize conservation programs more effectively than the untreated population. The results from the regressions in equation 2.5 for each of the four specifications of water conservation programs are presented in Table 2.12.[23]

In the pooled specification, participating in a conservation program reduces water use by almost 7%. Home surveys and toilet rebates each lead to a drop in consumption of over 10%. The clothes washer program is not statistically significant on its own, but it is jointly significant with the treatment interaction term at the 5% level. None of the interaction terms are statistically significant, but there is still valuable information in the sign and size of the coefficients.[24] Home surveys are essentially an additional information treatment by providing households with tailored advice to reduce consumption. The coefficient on the home survey interaction is positive but insignificant. It is difficult to posit why an interaction of two

[23]While there is not a sufficient sample size post-treatment to report the regression results in this paper, the conservation programs have a similar effect on water demand in Aquaville and Hydroburg. Tables presenting those results are available upon request.

[24]The lack of significance on the individual treatment programs is likely due to the sparseness of post-treatment participation data for each of the individual programs, whereas the coefficient on all programs is combining positive and negative effects.

Table 2.12: Interactions with Conservation Programs - Watertown

	Pooled		Rebates	
	(1)	(2)	(3)	(4)
	All Programs	Home Survey	Toilet	Clothes washer
Treatment Effect	-0.0501***	-0.0500***	-0.0530***	-0.0534***
	(0.0147)	(0.0147)	(0.0146)	(0.0147)
All Programs	-0.0674***			
	(0.0192)			
Treat*All	0.0128			
	(0.0294)			
Home Survey		-0.1274***		
		(0.0453)		
Treat*Survey		0.0656		
		(0.0505)		
Toilet Rebate			-0.1146***	
			(0.0382)	
Treat*Toilet			-0.0841	
			(0.0768)	
Clothes Washer Rebate				-0.0534
				(0.0351)
Treat*Clothes Washer				-0.0524
				(0.0639)
Household FEs	Yes	Yes	Yes	Yes
Year-Period FEs	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes
Adjusted R^2	0.158	0.158	0.158	0.157
Households	2,575	2,575	2,575	2,575
Observations	44,763	44,763	44,763	44,763

Notes: The dependent variable is the natural logarithm of water consumption in gallons per day. Robust standard errors clustered at the household level are reported in parentheses. Each regression equation includes a dummy variable if a household has participated in the conservation program by time t, as well as an interaction of the conservation variable with the treatment effect. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

information treatments would produce savings greater than the sum of the parts. The sign of coefficients on toilet and clothes washer rebates is congruent with the hypothesis that social

norms attract households to appliance rebates that otherwise would not have made those investments. However, the noise in the estimates require caution when making the argument that social norms alleviate concerns over additionality in rebate programs. Examining the results in Tables 2.11 and 2.12 together also demonstrates one of the methods through which social norm programs cause reductions in water consumption - by increasing participation rates in conservation programs. A back of the envelope calculation based on the coefficients in column (1) of Tables 2.11 and 2.12 suggests that approximately 20% of the ATE is due to participation in additional utility conservation programs. This is a strain of research that is important to pursue as data arrive from the other pilots. The interactions with existing programs may contribute to the overall ATE in a utility and could help describe the variation in ATE across pilots.

2.6 Future Work

This research is a portion of the broader project on the role of information in municipal water campaigns. WaterSmart is working with several other utilities that provide new opportunities for research. Several of the utilities do not want to implement randomized field experiments in the hopes of maximizing conservation gains immediately by targeting specific problem populations or by providing universal access to the normative message. Analyzing the data for utilities targeting subsets of their customer base re-introduces selection concerns and the identification of suitable control observations. This will be left for future research. An ongoing partnership with WaterSmart also provides opportunities to set up randomized experiments testing alternative messages across subgroups of those new pilots providing universal access. This has the same attractive features of WaterSmart's first three programs analyzed above by avoiding concerns of endogenous sample selection. The results of this paper suggest that there may be conservation gains available by targeting consumers to treatment based on their observables. Conducting a natural field experiment will test if additional conservation gains can be achieved by targeting the message based on observables.

There may be important interactions with water rates and normative messaging. These are difficult to compare across utilities because there are many unobserved shocks that vary across utility. Randomizing subgroups within a utility to exploit interactions between the social norms and water rates is another promising area of research. We hope to explore several of these avenues in future research.

2.7 Conclusion

Behavioral economic is a particularly attractive policy instrument to address environmental externalities where using Pigouvian taxes is politically intractable. Municipal water demand is a domain that must address issues of scarcity associated with rising populations, the exhaustion of cheap supply sources, and stronger demand for water devoted for ecosystem services. Climate change is expected to increase the variability of supply and increase the probability of extreme droughts. Zero profit constraints, concerns over equity, and the notion of access to water as a basic human right, limit the ability for public utilities to manage demand through the price mechanism. Utilities therefore seek a plethora of non-pecuniary demand side management practices. Among the suite of options, information campaigns using social norms are gaining momentum as an effective conservation instrument.

We analyze the data from three randomized field experiments that harness social norms as a water conservation tool. There is significant heterogeneity in the treatment effects both within and across pilots. While there are many dimensions that may impact response to treatment, we focus on the distribution of water use, environmental ideology, and housing values as observable drivers of heterogeneity. At an aggregate level Watertown, a small utility with a moderate climate and low baseline water use experiences the largest ATE of 5.34%. This is statistically different than the ATE for Hydroburg, with the smallest ATE of 1.7%. Interacting the treatment with quintiles of the baseline consumption corroborates the results of other social norm campaigns that treatment effects are largest for households with high baseline consumption.

Several results augment and address gaps in the prior literature. Contrary to Ferraro and Miranda (2013) who find a higher CATE for houses with assessed values above the median; the only significant heterogeneity in terms of housing values that we observe shows that higher housing values is correlated with lower savings. Across utilities there is monotonically decreasing relationship between average income and the ATE. Though it is difficult to draw inference from the three aggregate results in our data, this does provide evidence that social norms are effective with low-income households. We also observe two effects of ideology that work in opposite directions. In Aquaville stronger environmental preferences increase the CATE while in Hydroburg they correspond with smaller CATEs. In a series of unconditional quantile regressions we find that Watertown exhibits a QTE larger than the median at both the low and high end of the consumption distribution. Prior research (Allcott, 2011b; Ferraro and Miranda, 2013) finds a monotonic relationship between consumption quantiles and the QTE. This opens the possibility that social norms can be effective at reducing consumption for households that are already using water efficiently. These households may be more conscious of the financial and/or environmental costs associated with water consumption.

In addressing the mechanisms of how WaterSmart's HWR translate into water savings we look at the time pattern of individual-period treatment effects. We find little correlation of the treatment effect with weather variables that increase outdoor water use, suggesting that the HWRs do not have a large impact on outdoor use. Rather, the durability of savings over time is consistent with indoor water efficiency investments or systematic behavioral changes. Additionally, the time pattern of water savings varies across utilities, which indicates that demographics, climate, and utility policies affect the actions prompted by the normative message. One mechanism that we do identify is increased participation rates in additional utility conservation programs. This contributes up to 20% of the ATE in Watertown and demonstrates that social norms act as a complement rather than a substitute to existing conservation programs. To address water scarcity in the presence of regulatory and political pricing constraints utilities are seeking new policy instruments to reduce demand. Due to

the heterogeneity in treatment response, both within and across pilots, utilities must address underlying interactions in order to maximize the effectiveness of social norms as a water conservation instrument.

BIBLIOGRAPHY

Allcott, Hunt, "Consumers' Perceptions and Misperceptions of Energy Costs," *American Economic Review*, 2011, *101* (3), 98–104.

_ , "Social norms and energy conservation," *Journal of Public Economics*, 2011, *95* (9-10), 1082–1095.

_ **and Michael Greenstone**, "American Economic Association Is There an Energy Efficiency Gap?," *American Economic Review*, 2012, *26* (1), 3–28.

_ **and Sendhil Mullainathan**, "Behavioral Science and Energy Policy," *Science*, 2010, *327* (6970), 1204–1205.

_ **and Todd Rogers**, "How Long Do Treatment Effects Last? Persistence and Durability of a Descriptive Norms Intervention's Effect on Energy Conservation," 2012.

_ **and** _ , "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," 2012.

Benartzi, Shlomo and Richard H Thaler, "Heuristics and Biases in Retirement Savings Behavior," *Journal of Economic Perpectives*, 2007, *21* (3), 81–104.

Bennear, Lori S, Jonathan M Lee, and Laura O Taylor, "Municipal Rebate Programs for Environmental Retrofits : An Evaluation of Additionality and Cost-Effectiveness," *Journal of Policy Analysis and Management*, 2013, *32* (2), 350–372.

Bitler, Marianne P, Jonah B Gelbach, and Hilary W Hoynes, "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments," *American Economic Review*, September 2006, *96* (4), 988–1012.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes, "Distributional impacts of the Self-Sufficiency Project," *Journal of Public Economics*, April 2008, *92* (3-4), 748–765.

Bondt, Werner FM De and Anil K Makhija, "Throwing good money after bad?: Nuclear power plant investment decisions and the relevance of sunk costs," *Journal of Economic Behavior & Organization*, 1988, *10* (2), 173–199.

Borah, Bijan J and Anirban Basu, "Highlighting Differences Between Conditional and Unconditional Quantile Regression Approaches Through an Applicatino to Assess Medication Adherence," *Health economics*, 2013, *22*, 1052–1070.

Braden, John B, Xia Feng, and DooHwan Won, "Waste sites and property values: a meta-analysis," *Environmental and Resource Economics*, 2011, *50* (2), 175–201.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics: Methods and Applications*, Cambridge University Press, 2005.

Clark, D. E. and T. Allison, "Spent nuclear fuel and residential property values: The influence of proximity, visual cues, and public information," *Papers in Regional Science*, 1999, *78*, 403–421.

_ , L. Michelbrink, T. Allison, and W. C. Metz, "Nuclear power plants and residential housing prices," *Growth and Change*, 1997, *28*, 496–519.

Costa, Dora L. and Matthew E. Kahn, "Energy Conservation "Nudges" and Environmentalist Ideology: Evidence From a Randomized Residential Electricity Field Experiment," *Journal of the European Economic Association*, June 2013, *11* (3), 680–702.

Crump, Richard, V Joseph Hotz, Guido Imbens, and Oscar Mitnik, "Moving the goalposts: Addressing limited overlap in the estimation of average treatment effects by changing the estimand," 2006.

Dalhuisen, Jasper M., Raymond J. G. M. Florax, Henri L. F. de Groot, and Peter Nijkamp, "Price and Income Elasticities of Residential Water Demand: A Meta-Analysis," *Land Economics*, May 2003, *79* (2), 292.

Davis, Lucas W, "The effect of power plants on local housing values and rents," *Review of Economics and Statistics*, 2011, *93* (4), 1391–1402.

Djebbari, Habiba and Jeffrey Smith, "Heterogeneous impacts in PROGRESA," *Journal of Econometrics*, July 2008, *145* (1-2), 64–80.

Dolan, Paul and Robert Metcalfe, "Better neighbors and basic knowledge : a field experiment on the role of non-pecuniary incentives on energy consumption ," 2011.

Downs, Julie S, George Loewenstein, and Jessica Wisdom, "Strategies for Promoting Healthier Food Choices," *American Economic Review: Papers & Proceedings*, 2009, *99* (2), 159–164.

Elster, Jon, "Social Norms and Economic Theory," *Journal of Economic Perspectives*, 1989, *3* (4), 99–117.

Espey, M., J. Espey, and W. D. Shaw, "Price elsasticity of residential demand for water: A meta-analysis," *Water Resources Research*, 1997, *33* (6), 1369–1374.

Ferraro, By Paul J, Juan Jose Miranda, and Michael K Price, "The Persistence of Treatment Effects with Norm-Based Policy Instruments : Evidence from a Randomized Environmental Policy Experiment," *American Economic Review*, 2011, *101* (3), 318–322.

Ferraro, Paul J. and Juan José Miranda, "Heterogeneous Treatment Effects and Mechanisms in Information-Based Environmental Policies: Evidence from a Large-Scale Field Experiment," *Resource and Energy Economics*, April 2013, *Forthcomin.*

Ferraro, PJ and MK Price, "Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment," *Review of Economics and Statistics*, 2013, *95* (1), 64–73.

Fipro, Sergio, "Efficient Semiparametric Estimation of Quantile Treatment Effects," *Econometrica*, 2007, *75* (1), 259–276.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux, "Unconditional Quantile Regressions," *Econometrica*, 2009, *77* (3), 953–973.

Folland, Sherman and Robbin Hough, "Externalities of nuclear power plants: further evidence," *Journal of Regional Science*, 2000, *40* (4), 735–753.

Frölich, M and B Melly, "Quantile treatment effects in the regression discontinuity design," *Journal of Econometrics*, 2012, *168*, 382–395.

Frölich, Markus and Blaise Melly, "Unconditional Quantile Treatment Effects Under Endogeneity," *Journal of Business & Economic Statistics*, July 2013, *31* (3), 346–357.

Galster, George C, "Nuclear power plants and residential property values: a comment on short-run vs. long-run considerations," *Journal of Regional Science*, 1986, *26* (4), 803–805.

Gamble, Hays B and Roger H Downing, "Effects of nuclear power plants on residential property values," *Journal of Regional Science*, 1982, *22* (4), 457–478.

Gaudin, S., "Effect of price information on residential water demand," *Applied Economics*, March 2006, *38* (4), 383–393.

Gawande, Kishore and Hank Jenkins-Smith, "Nuclear waste transport and residential property values: estimating the effects of perceived risks," *Journal of Environmental Economics and Management*, 2001, *42* (2), 207–233.

Gneezy, Uri and John a List, "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments," *Econometrica*, September 2006, *74* (5), 1365–1384.

Grafton, R. Quentin and Michael B. Ward, "Prices versus Rationing: Marshallian Surplus and Mandatory Water Restrictions*," *Economic Record*, September 2008, *84* (September 2008), S57–S65.

Halvorsen, R. and R. Palmquist, "The interpretation of dummy variables in semilogarithmic equations," *American Economic Review*, 1980, *70* (3), 474–75.

Harrison, Glenn W and John A List, "Field Experiments," *Journal of Economic Literature*, 2004, *42* (4), 1009–1055.

Heintzelman, Martin D and Carrie M Tuttle, "Values in the wind: A hedonic analysis of wind power facilities," *Land Economics*, 2012, *88* (3), 571–588.

Hofert, Marius and Mario V Wüthrich, "Statistical Review of Nuclear Power Accidents," *Asia-Pacific Journal of Risk and Insurance*, 2011, *7* (1).

Ito, Koichiro, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," 2012.

_ , "How Do Consumers Respond to Nonlinear Pricing? Evidence from Household Water Demand," 2013.

Johnson, Eric J and Daniel Goldstein, "Do Defaults Save Lives ?," *Science*, 2003, *302*, 1338–1339.

Kiel, Katherine A and Katherine T McClain, "House prices during siting decision stages: the case of an incinerator from rumor through operation," *Journal of Environmental Economics and Management*, 1995, *28* (2), 241–255.

Koenker, Roger and Gilbert Basset, "Regression Quantiles," *Econometrica*, 1978, *46* (1), 33–50.

Kuethe, Todd H and Roman Keeney, "Environmental Externalities and Residential Property Values: Externalized Costs along the House Price Distribution," *Land economics*, 2012, *88* (2), 241–250.

Landry, Craig E., Andreas Lange, John A. List, Michael K. Price, and Nicholas G. Rupp, "Toward an understanding of the economics of charity: Evidence from a field experiment," *Quarterly Journal of Economics Journal of Economics*, 2006, (May).

Levitt, Steven D. and John a. List, "Field experiments in economics: The past, the present, and the future," *European Economic Review*, January 2009, *53* (1), 1–18.

List, JA and D Lucking-Reiley, "Demand reduction in multiunit auctions: Evidence from a sportscard field experiment," *The American Economic Review*, 2000, *90* (4), 961–972.

Mansur, Erin T. and Sheila M. Olmstead, "The value of scarce water: Measuring the inefficiency of municipal regulations," *Journal of Urban Economics*, May 2012, *71* (3), 332–346.

Metz, William C, "Potential negative impacts of nuclear activities on local economies: rethinking the issue," *Risk Analysis*, 1994, *14* (5), 763–770.

Nataraj, Shanthi and W. Michael Hanemann, "Does marginal price matter? A regression discontinuity approach to estimating water demand," *Journal of Environmental Economics and Management*, March 2011, *61* (2), 198–212.

Nelson, Jon P, "Three Mile Island and residential property values: Empirical analysis and policy implications," *Land Economics*, 1981, pp. 363–372.

Powell, David, "A New Framework for Estimation of Quantile Treatment Effects: Non-seperatble Disturbance in the Presence of Covariates," 2013.

Renwick, Mary E. and Richard D. Green, "Do Residential Water Demand Side Management Policies Measure Up? An Analysis of Eight California Water Agencies," *Journal of Environmental Economics and Management*, July 2000, *40* (1), 37–55.

Renwick, Mary E and Sandra O Archibald, "Demand Side Management Policies for Residential Water Use: Who Bears the Conservation Burden?," *Land Economics*, 1998, *74* (3), 343–359.

Semenza, Jan C., Carol H. Rubin, Kenneth H. Falter, Joel D. Selanikio, Dana W. Flanders, Holly L. Howe, and John L. Wilhelm, "HEAT-RELATED DEATHS DURING THE JULY 1995 HEAT WAVE IN CHICAGO," *The New England Journal of Medicine*, 1996, *335*, 84–90.

Sexton, Steven E, "Energy Price Salience and Residential Energy," 2010.

Shin, Jeong-Shik, "Perception of Price When Price Information is Costly: Evidence from Residential Electricity Demand," *The Review of Economics and Statistics*, 1985, *67* (4), 591–598.

Thaler, Richard H and Cass R Sunstein, "Libertarian Paternalism," *American Economic Review: Papers & Proceedings*, 2003, *93* (2).

Woolridge, Jefferey M., *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press, 2002.

Worthington, Andrew C. and Mark Hoffman, "an Empirical Survey of Residential Water Demand Modelling," *Journal of Economic Surveys*, July 2008, *22* (5), 842–871.

Appendix A

APPENDIX FOR CHAPTER 1

A.1 Challenges to US Nuclear Energy and the Federal Strategy

Nuclear power as a source of clean energy remains near the forefront of US energy policy goals. However, a study by the Princeton Carbon Management Initiative estimated that for nuclear power to play a significant role in cutting greenhouse emissions, the industrys capacity would have to triple worldwide over the next 50 yearsa rate of 20 new large reactors a year. Most of the 40% increase in nuclear power production since the 1990s[1], the equivalent of 29 new 1,000 MWe reactors, has not been achieved through increased capacity, but from a better ability to yield more energy from the capacity that currently exists. Plants went from a capacity factor (the percentage of power a plant can produce if it ran continuously) of 56% in 1980 and 66% in 1990 to a level of around 90% today[2].

With the capacity factor close to peak efficiency, increases in nuclear power generation must be achieved through new reactors at new or existing plants. The early days of nuclear power were full of optimism over the energy of the future. The phrase too cheap to meter coined by Lewis Strauss, the Chairman of the United States Atomic Energy Commission in 1954, repeatedly resurfaced to advocate for commercial nuclear power. However, the reality is much different. To construct a new plant, energy companies face massive, unpredictable, multiyear (often decades long) construction projects and years of wading through licensure and regulatory procedures. The thirty-five nuclear power projects still under construction in January 1985 were estimated to cost six to eight times the amount originally budgeted.De Bondt and Makhija (1988).

Even the addition of a new reactor to existing plants is often unpredictably costly. The Tennessee Valley Authority is expected to complete its second reactor at Watts Bar in 2015, years later than expected and at a cost to utility consumers of over $4 billion, more than twice the expectation set in 2007.

---

[1]World Nuclear Association, http:www.world-nuclear.org/info/inf41.html, Nuclear Power in the USA

[2]Nuclear Energy Institute, http://www.nei.org/Knowledge-Center/Nuclear-Statistics/US-Nuclear-Power-Plants

In US history, at least 15 plants have been partially constructed at great expense without ever operating. Before 1998, 12 plants were closed after initial operation due to safety concerns or operating costs so high that it was cheaper to buy power from other sources. Two plants, the San Onofre Nuclear Power Plant in California and the Crystal River Plant in Florida, were permanently shut down during the first half of 2013 because of structural damage too expensive to fix, the first of possibly many plants in the aged US nuclear fleet to cancel operations. The energy company Dominion closed the doors on the Kewaunee plant in May, 2013 citing an inability to compete with natural gas prices in the unregulated utility markets of Wisconsin. Entergy, the company who owns the Vermont Yankee power plant, followed suit and announced plans to close in the last quarter of 2014.

Consistently low natural gas prices since 2009 will continue to plague nuclear power in unregulated states. In 2010, an executive at Exelon, the largest nuclear energy producer in the US, was quoted saying that natural gas prices would have to reach $8 per 1 million BTU and a carbon fee of $25 a ton to make a nuclear project worthwhile[3]. During the first 13 days of November 2013, natural gas was only $3.47 per 1 million BTU[4].

Even in these conditions, there are 28 license applications[5] to build new nuclear reactors pending. Five reactors are currently under construction: two at Georgia's Vogtle plant, two at the Virgil C. Summer plant in South Carolina, and the fifth at Watts Bar in Tennessee. The plants in Georgia and South Carolina are owned and operated by regulated utilities and so insulated from the price of natural gas; the Energy Policy Act of 2005 is necessary to the success of other proposals in unregulated states.

The Energy Policy Act of 2005, easily approved by the US legislature, allocated billions of dollars to support increased nuclear energy capacity. It provides production tax credits on the first 6,000 MWe of new nuclear capacity during the first eight years of operation at the same rate available to wind power production, 2.1 cents/kWh. It also provides federal

---

[3]Matthew Wald. Sluggish Economy Curtails Prospects for Building Nuclear Reactors, October, 10, 2010

[4] http://www.eia.gov/naturalgas/weekly/

[5] http://www.nrc.gov/reactors/new-reactors/col.html, updated Sep23, 2013

risk insurance of $2 billion in the case of regulatory delays and up to $18.5 billion towards federal loan guarantees for the construction of new nuclear reactors. The White House 2014 fiscal budget allocates $735 million to the Office of Nuclear Energy, which includes funding for the R&D of advanced small modular reactors, the supposed next generation of nuclear power.[6]

A.2 Surveys and Stated Choice Studies: the Perception of Nuclear Power

Very few studies have used stated preference techniques to reveal opinions about nuclear energy. To our knowledge, there are only two examples that focus on accidents at nuclear plants specifically and a few more that focus on the welfare loss from nuclear-waste transport.

Alvarez-Farizo and Howard (2009) formed two community workshops, one in the Cumbria (United Kingdom) and the other in Zaragoza (Spain), to focus on the social impacts of clean-up procedures in the case of a hypothetical nuclear accident. Though the willingness to pay to correct specific damages was determined, the authors do not focus or report these numbers, and rather focus on the methodologies used to infer community preferences and priorities.

Schneider and Zweifel (2004) use a stated choice methodology to illicit the willingness to pay of Swiss citizens for relief from the financial consequences of a severe nuclear accident. They asked a sample of Swiss citizens to choose between a status quo and an alternative, which was varied to describe different types of power sources and the degree of liability insurance. They found that for a power type reflecting nuclear power, described by having few black outs, unresolved problems with waste disposal, a maximum possible loss of US$130 billion (Schneider and Zweifel, 2004), and initial liability insurance coverage of only 1 percent, a Swiss citizen with median characteristics is willing to pay 1.3 percent more for electricity to have an additional 1 percent of coverage. have found that these stated preferences however do not match up with the revealed preferences. In this paper we focus on the revealed preference data only.

---

[6] http://www.world-nuclear.org/info/Country-Profiles/Countries-T-Z/USA–Nuclear-Power-Policy/

Riddel (2011) and Riddel and Shaw (2006) use a survey of Nevada residents along the proposed nuclear-waste transportation route to the proposed national repository in Yucca Mountain. The papers differ predominantly in their modeling specifications and are both meant to contribute novel techniques in the modeling of uncertain subjective risk to the risk literature as opposed to evaluating the willingness to accept nuclear-waste transport. However, both papers report such estimates. Using survey respondents answers to questions concerning their willingness to accept compensation to remain along the transportation route or move away. Riddel (2011) and Riddel and Shaw (2006) find that high-level nuclear-waste transport amounts to an annual welfare loss of $6,583 and $5,400 per household, respectively.

In contrast, Gawande and Jenkins-Smith (2001) use revealed preference techniques to estimate a premium associated with being 5 miles away from the nuclear waste shipments through Charleston County in California. This premium is estimated to be 3 percent of home values. With a sample median home value of $66,710 this amounts to only $2,001.

A few assumptions and a quick back of the envelope calculation compares the vastly different results of Gawande and Jenkins-Smith (2001) using revealed preference techniques to the results of Riddel and Shaw (2006) using stated preferences. Using the ratio of mean household incomes reported in the two papers ($51,100 and $33,030, respectively), we assume that the Riddel and Shaw (2006) estimates from Clark County, NV, where the willingness to accept the risk from nuclear transport is valued at $5,400 annually, would have amounted to $3,490 annually in Charleston County, CA. Assuming a home tenure of ten years and a discount rate of 6%, this would translate to a willingness to accept of $25,690. This value is extreme in comparison to the $2,001 home value premium reported by Gawande and Jenkins-Smith (2001).

In the light of the controversial role that nuclear power has played over its lifetime, it maybe unsurprising that stated preference and revealed preference estimates concerning nuclear power are so different. Metz (1994) reviews the paradoxes apparent in survey results focused on nuclear power plants, nuclear waste transport, and waste storage. For example,

Metz (1994) notes that 80% of the US population lives within 100 miles of a nuclear power plant, yet 26 and 20 percent of Las Vegas residents say they would be unwilling to live within 200 and 500 miles, respectively.

A.3 Discussion of Hedonic Method in the Current Context

The hedonic method is commonly used to value amenities for which there are no formal markets. It does this by revealing the implicit prices of attributes that compose differentiated products. Rosen (1974) originally modeled competitive housing markets with full information to buyers and sellers. He assumed a house could be fully described by a vector of observable attributes $\mathbf{x} = (x_1, , x_K)$ including both structural and community characteristics.

The hedonic price schedule, $P(\mathbf{x})$ is the equilibrium result of buyers and sellers optimization problems. For the purposes of this paper, we focus on the optimizing behavior of buyers. Each period, individuals are assumed to maximize utility subject to a budget constraint:

$$\max_{\mathbf{x},c} \quad U(x, c; \delta) \quad \text{s.t.} \quad Y = P(x) + c \tag{A.1}$$

where: c is a composite consumption numeraire good and δ is a preference parameter specific to the individual.

The first order conditions of this optimization problem lead to the relationship that the marginal rate of substitution between the $\mathbf{x}_k^{th}$ attribute and the consumption good equals the slope of the hedonic price function with respect to z_k:

$$\frac{\frac{\partial U(\mathbf{x}, Y - P(\mathbf{x}); \delta)}{\partial z_k}}{\frac{\partial U(\mathbf{x}, Y - P(\mathbf{x}); \delta)}{\partial c}} = \frac{\partial U(\mathbf{x}, Y - P(\mathbf{x}); \delta)}{\partial z_k} \tag{A.2}$$

An individual consumers maximum bid function $\theta(\mathbf{x}; U = u_0, Y)$, drawn concave from below in\$, x_k space, can be interpreted as the consumers indifference curve between c and x_k. As the value of the appealing attribute increases along the bid function, the consumer is willing to give up the consumption numeraire and still maintain utility levels. With the

house going to the highest bidder in competitive equilibrium, a given individuals maximum bid function is tangent to the hedonic price gradient. Then, the hedonic price function can be shown to be an envelope of consumers maximum bid functions. Therefore, the empirical estimation of the hedonic price function reveals consumers marginal willingness to pay for incremental increases in the amenity.

Interpreting the x_k^{th} attribute as proximity to nuclear power plants serves as our proxy for increased risk of a nuclear accident and simultaneous increased access to new jobs. Nuclear power plants may also increase the quality of the area, either through their contribution to the tax base or from their own public works to compensate for perceived dangers and unsightliness of the plant. For instance, some power companies have created parks in the open spaces around plants. We would expect individuals with greater preference for safer distances to locate at higher values of x_k along the hedonic price function, whereas those attracted to high-tech job availability and other positive amenities will move closer. This introduces modeling issues as individuals with greater accident risk aversion determine the slope at greater distances and job seeking individuals with low risk aversion determine the slope of the hedonic price function at lower distances. This self-sorting makes it difficult to generalize estimates for typical populations.

The hedonic method was originally conceived to model the behavior of individuals. Unfortunately, a reasonable dataset with individual housing level data dating back to 1970 is unavailable. The dataset used in this paper contains values aggregated at the census tract level. Aggregating the hedonic price function by census tract and estimating on a national scale may introduce problems (Chay & Greenstone, 2005). It requires we assume a national U.S. housing market. This means consumers have the option and ability to select their optimal housing out of the entire U.S. sample. Aggregating the hedonic model construed by Rosen at the individual consumer level could introduce bias by obscuring individual level non-linear relationships between nuclear proximity and housing values. Chay and Greenstone test for bias by performing cross-sectional regressions with their tract level data. They

then match the model specifications and locations studied to previously published literature using individual level data. In comparing estimates, they conclude that simply aggregating is not a significant source of bias.

We avoid this altogether by defining proximity dichotomously. For the purposes of this paper, we focus on the empirical estimation of the hedonic price premium, or discount, associated with proximity to nuclear power plants to reveal the willingness of individuals to pay for a dichotomous improvement, close or not close.

A.4 Identification

A.4.1 Defining Treatment and Geographic Sample Exclusions

Defining treatment as percent of area overlapped by buffers is informed by the great variation in tract size and shape. Tracts were originally defined to surround similar populations of 2,500 to 8,000 people. Within 5 miles of city centers, tracts can be quite small, averaging 1.01 square miles. However, tracts within 10-mile buffer rings (Plume Exposure Pathways) around nuclear power plants in our sample average 5.96 square miles. This, compounded with the irregular shaping of census tracts, led us to conclude that using distances to tract centroids would be less informative.

The following rules, each with a parameter (in bold) whose value is left to the judgment of the researcher, define the treatment groups and baseline sample exclusions.

Defining treatment:

- Plume exposure pathways (10-mile buffers) cover over 80% of census tract area.
- 5-mile buffers cover over 50% of census tract area.
- Exclude untreated tracts within 20 miles of the nuclear power plant sets.

Defining exclusions to eliminate confounding influence:

- Exclude untreated tracts within 20 miles of nuclear power plants not in the target nuclear power plant set.

In the pursuit of the best parameter values to use, we tested the sensitivity of the treatment groups descriptive statistics and sample size to various specifications. For example, when defining treatment such that the plume exposure pathway covers at least 50% of census tracts, the means of important covariates were very sensitive to the values chosen for the exclusion parameters. When investigating these treated areas geographically we noticed that around some plants the treated area was spatially irregular and wide spread, whereas around others it was orderly and more or less confined to the plume exposure pathways. This irregularity is primarily due to the large variation in census tract area. The specification weve settled on, to require that 80% of the tracts area be covered, leads to greater consistency around plants. With this specification, the effect of augmenting the exclusion buffer out to 30 miles has little effect on descriptive statistics and so weve opted for a 20-mile exclusion to retain more control observations. In all, the selection of these parameters was an exercise in good judgment and justifiability.

A.4.2 Endogeneity due to Placement

For the identification of the effect of nuclear power plants on local markets, the presence of a nuclear power plant would ideally be random. Unfortunately, and perhaps obviously, this is not the case. Commercial nuclear power sites are determined through a long regulatory process often riddled with public backlash and civic court proceedings. Without a random treatment, we achieve balance between treatment and groups through subsampling the set of potential control observations[7] and propensity score weighting. The propensity score literature emphasizes that only those variables that simultaneously predict treatment and

[7]Limiting the control sample before weighting reduces the variance of the final estimators (Crump, Hotz, Imbens, and Mitnik, 2007) by first eliminating observations incomparable to the treated sample.

the intended dependent variable need be included in the estimation of propensity score weights. We investigated nuclear power plant siting policy to inform the set of variables used to predict treatment.

Nuclear Siting Policy and Estimating Important Variables in the Identification of Control Groups

The Atomic Energy Act of 1954 broke the tight government monopoly over nuclear technology and opened the way for commercial uses[8] of nuclear materials to be regulated and promoted by the newly created Atomic Energy Commission (AEC). The AEC held a mandate to (1) continue the production and testing of nuclear weapons; (2) promote private use of atomic energy for peaceful applications; and (3) ensure the safety of commercial nuclear power (Walker, 1992). Despite strong efforts to fulfill the second mandate, by 1955 only 5 power companies announced an intention to build small nuclear facilities (Walker). Desperate to compete with the U.S.S.R. for international prestige, demonstrate the peaceful potential of nuclear energy after the catastrophic damage to Japan, and mitigate dependence on fossil fuels, the AEC stepped up its commitment to the second mandate, the promotion of commercial nuclear energy.

By 1966 the market for commercial energy was booming with 20 nuclear units purchased that year, 31 in 1967, and 17 in 1968. Such a dramatic increase in caseload caused problems for the AEC charged with both the promotion of nuclear energy and the regulation of an industry so new that what constituted safe and consistent operation of commercial nuclear reactor was still being determined.

Aside from caseload, the AEC was increasingly confronted with public pressure against their conflicting mandates. The AEC was abolished in 1974 and replaced by the Nuclear Regulatory Commission (NRC), which unlike the AEC had no promotional duties.

---

[8]While private firms could own the plants and equipment, they had to lease nuclear fuel from the U.S. government, as private ownership of nuclear materials did not become possible until 1964 (Walker 1992, pg 23).

This overburdened, complex, and controversial system managed the licensing and initial inspection of most of the nuclear power plants operating today. It was in this policy environment that the major issues of nuclear plant siting were debated in a create-as-you-go regulatory process. The major issues to consider in nuclear plant siting include: adequate access to water for cooling and maintaining a safe distance from metropolitan and seismically active areas. The plants placed in the 1970s achieve this goal to varying degrees.

Distance to City Centers

Utility companies, in trying to avoid the high cost of energy transmission to consumers, often pursued sites close to highly populated areas. However, the newness of nuclear technology, the untried reliability of safety features by 1970, and public outcry made the regulatory agencies more likely to approve plants located at safer distances. To illuminate the ability of proposed plants to meet safety requirements, an applicant for a commercial nuclear power construction permit needed to define the population center distance of the proposed site as the distance from a reactor to the nearest boundary of a populated center amounting to 25,000 residents[9]. Distance to major city centers is also influential in predicting housing values, household income, and many other dependent variables of interest to our investigation.

We use geographical information systems to estimate the distance from tract centroids to cities with a population greater than 25,000 in each of our control years: 1970 to estimate the effect of plant placement; 1980 to estimate the effect of the Chernobyl disaster or plant closures during the 1980s; and 1990 to estimate the effect of plant closures during the 1990s. The population of major cities in 1990 is readily available, as is county populations in 1970, 1980, and 1990. To estimate the population of city centers in 1970 or 1980, we use the county population growth rate from the base year to 1990 and project backwards to create a localized Population Adjustment Factor. For the counties whose boundaries changed between 1970 or 1980 and 1990, weve corrected for the changing area of a county over the decades by first

[9]Site regulations did not specify precise distance requirements, allowing licensees to meet approval through engineered safety measures as well as distance (Walker 1992, pg 60-62).

calculating the base years population normalized to county areas in 1990. [10]

Distance to Adequate Water Sources

Nuclear power plants require steady and predictable access to cold water sources for multiple purposes: to generate steam to drive turbines; to condense steam and recycle it within the system; and to maintain safe temperatures in the prevention of accidents. Indeed, most commercial nuclear accidents have been caused by the malfunction of the cooling systems. Water distance can also serve as a proxy for proximity to recreation areas and environmental amenities. The precise definition of what constitutes a sufficiently large water source is unavailable. We noted, however, that 90% of commercial nuclear power plants are placed within 5 miles of water sources with an area greater than three square miles. We then use distance from tract centroid to water bodies of this size or larger in our estimations.

Seismic Risk

In the 1960s and 1970s as nuclear siting policy was being formed, little was understood about seismic risk or even the identification of fault lines. For instance, long civil suits and appeals over apparent (or not so apparent) fissures held up the Bodega Bay plant until it was eventually canceled in the mid-1970s. For our estimations, we pursued the inclusion of a variable capturing the distance from tract centroids to the nearest earthquake epicenter from the 100 years preceding 1970. This variable had little impact on the estimation of propensity score weights and is arguably unrelated to many of our dependent variables of interest and so was excluded in our final analysis.

Estimation of Propensity Score Weights

Spatial proximity variables such as distance to city centers with a population greater than 25,000 in the control year, in this case 1970, are significant in predicting treatment. Other consistently significant predictors are the percent of the population who is Hispanic, average housing values, the unemployment rate, the tenure rate, and the portion of residential

[10]Sources: City Shapefile with population in 1990 is from http://www.colorado.edu/geography/foote/maps/assign/cityrank
County populations and shape files in 1970 and 1990 are from www.NHGIS.org.

housing made up of multi-unit structures within census tracts.

Subsampling the set of control observations to hold only those covariate values observed within the treatment groups has a large effect on the predicted probabilities of treatment. An important effect on the predicted probabilities is the increased common support between the control and treatment groups. The following figures provide the distribution of the predicted probabilities of treatment for the control group (on the left) and the treatment group (on the right). Without subsampling, relative to the full set of control observations, the number of treated tracts is very small. In the case of rare events like these, logistic regression tends to under predict treatment. After subsampling, a greater portion of the resulting control sample, as well as the treatment group, is assigned higher probabilities of treatment.

To ensure all treatment observations have relevant control observations within the sample, and vise-versa, we drop all observations with a predicted probability of treatment beyond that observed in the opposite group. Thankfully, this is rarely necessary. Regressions run without ensuring common support in this analysis show insignificantly different results from those meeting this condition.

Appendix B

APPENDIX FOR CHAPTER 2

B.1 Pilots

WaterSmart partners with water utilities to send Home Water Reports, which provide information to the household on their water use relative to a similar cohort of households as well as personalized conservation recommendations. All of WaterSmart's current pilot programs are with California utilities; the names of these utilities have been disguised to protect the privacy of the water utilities' customers. Three of these programs have a randomized experimental design and have been active long enough to analyze for a combined initial sample of 805,467 total water use observations.[1] WaterSmart's first pilot, Watertown, has completed its experimental phase; all customers in this utility have since begun to receive home water reports. The experiments in the other two pilots are on going; water observations used in this paper for these pilots are as recent as August 2013. Sample sizes, the date of the first home water report sent, experiment end dates, and the number of households treated in each pilot can be found in Table B.1.

Table B.1: Initial Sample Sizes

Pilot	Start Date	End Date	N: Obs	N: Post-Treat	HHs	Treated HHs
Watertown	2011-09-20	2013-01-01	44,822	16,266	2,575	992
Aquaville	2012-06-28	on going	93,664	18,583	3,096	1,547
Hydroburg	2012-07-01	on going	666,981	115,649	11,710	1,184

B.2 Irregular Period Lengths and Staggered Reads

WaterSmart provides water meter reads for each household in their pilot program, though the data are originally generated from the water utilities themselves. The individual utilities define regularly spaced official read periods. For example, in Aquaville and Watertown, a

[1] This initial dataset is after very preliminary data cleaning. Observations before January 2008 have been dropped due to the limited number of households with water reads spanning back before that time. We've also dropped water read observations with extreme outliers: those beyond the median $\pm 28\times$interquartile range by pilot or those equal to zero.

new read period begins every two months from an initial date of January 1st. Hydroburg, however, begins a new official read period every month. With some exceptions, a water read is taken for each household during every read period. For individual households however, the time period between one read and then next, at which point the accumulation of water used since the last read is recorded, is irregular.

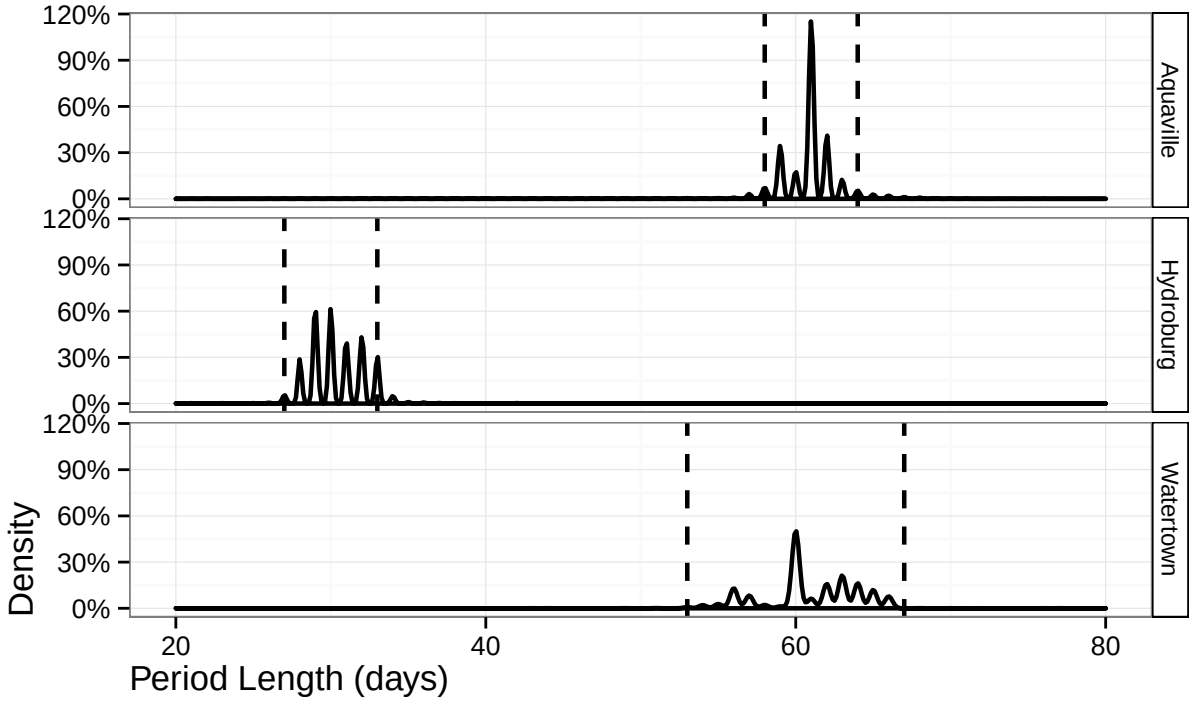
Individual read period lengths are not uniform across households, even within a given pilot. Table B.2 provides selected quantiles of period length by pilot. We determine what defines a short or long period relative to the distribution of period lengths within the pilot. Currently, the high end cutoff for each pilot is set by the median period length plus the inter-quantile range between the 85th and 15th percentile. The low end cutoff is the median less this inter-quantile range. The specific quantiles for these ranges were chosen by inspection of the distributions in Figure B.1. These cutoffs are given in Table B.3 below. We consider short read periods to be a more serious deviation from the norm and run versions of our regressions excluding them. We assume that a utility must have had a reason, such as a leaking pipe, to return after a short period to take another water read and restart the meter for the next period.

Table B.2: Period Length: high and low quantiles by pilot

	Watertown	Aquaville	Hydroburg
0th	12.0	1.0	1.0
1st	54.0	35.0	27.0
5th	56.0	58.0	28.0
25th	60.0	60.0	29.0
50th	60.0	61.0	30.0
75th	63.0	62.0	32.0
95th	66.0	64.0	33.0
99th	66.0	67.0	35.0
100th	155.0	143.0	50.0

Figure B.2 shows the distribution of water reads across the days of the month on which

Figure B.1: Period Length Distributions: Identifying short and long periods



Notes: The graph provides the distribution of water read period lengths. The vertical dashed lines indicate the cutoffs we use to define short or long read periods.

Table B.3: Short and long read periods relative to pilot

Pilot	Cutoff: long period	Cutoff: short period
Aquaville	64	58
Hydroburg	33	27
Watertown	67	53

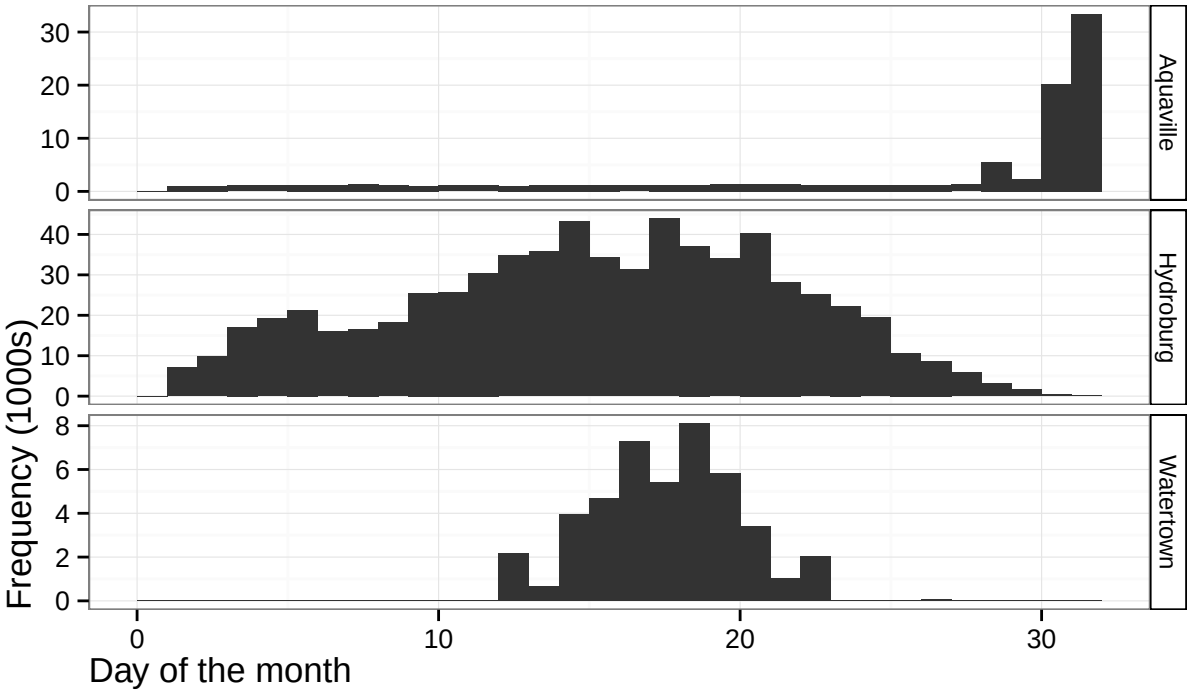
the meter was read by pilot. Notice that in Aquaville the density is much higher towards the end of the month, whereas in Hydroburg and Watertown meter reads are centered near the middle of the month. The staggered and irregular nature of bimonthly meter read periods leads to an even month preference in Watertown, see Table B.4. Only a handful of treatment and control observations are read during odd months. Due to the small number of these odd month meter reads, these observations are easily dropped. This avoids potential problems that may arise if the water reads pulled during odd months are systematically different from reads taken during even months.

Table B.4: Observation counts by month and pilot.

	Control Observations			Treated Observations		
Month	Aquaville	Hydroburg	Watertown	Aquaville	Hydroburg	Watertown
1	4,050	6,626	9	4,046	6,540	9
2	4,215	6,646	2,612	4,184	6,551	2,884
3	3,961	6,676	4	3,954	6,563	2
4	4,423	6,690	2,606	4,379	6,591	2,874
5	4,031	6,716	0	4,007	6,605	1
6	4,503	6,708	3,479	4,450	6,607	3,823
7	4,025	5,628	7	4,046	5,542	7
8	3,827	5,641	3,475	3,758	5,560	3,821
9	3,310	5,642	1	3,302	5,572	1
10	3,784	5,664	3,440	3,746	5,580	3,776
11	3,219	5,673	1	3,201	5,594	3
12	3,639	5,682	3,402	3,604	5,606	3,738

The time based fixed effects are important to control the effect of seasonality on water use.

Figure B.2: Frequency of water reads by pilot and day of the month



However, for pilots with bimonthly meter reads, if meters read on odd months are systematically different from water reads drawn during even months in a way that is correlated to the treatment group or other cross-sectional covariates of interest, then the year-month dummies will capture some of this cross-sectional variation and corrupt the regression coefficients.

Fortunately, we can avoid these difficulties by instead using utility-defined reading periods. Table B.5 shows that the number of readings are more balanced across reading periods.

Table B.5: Observation counts by reading period and pilot.

	Control Observations			Treated Observations		
Period	Aquaville	Hydroburg	Watertown	Aquaville	Hydroburg	Watertown
1	8,265	6,626	2,621	8,230	6,540	2,893
2	8,384	6,646	2,610	8,333	6,551	2,876
3	8,534	6,676	3,479	8,457	6,563	3,824
4	7,852	6,690	3,482	7,804	6,591	3,828
5	7,094	6,716	3,441	7,048	6,605	3,777
6	6,858	6,708	3,403	6,805	6,607	3,741
7	0	5,628	0	0	5,542	0
8	0	5,641	0	0	5,560	0
9	0	5,642	0	0	5,572	0
10	0	5,664	0	0	5,580	0
11	0	5,673	0	0	5,594	0
12	0	5,682	0	0	5,606	0

To further ensure balanced samples for panel regressions using year-period fixed effects, we must ensure that each household has only one water read per year-period to avoid overweighting the contribution of individual households during certain periods. Removing short periods will remove many duplicates, but not all. We randomly drop one water read in each household-year-period pair where there are duplicate observations. Table B.6 provides the number of additional observations dropped to balance the panel samples.

Table B.6: Ensuring a balanced panel: identify multiple obs within year-period.

Pilot	Total Obs	Duplicates	Duplicates (no short periods)
Aquaville	93,664	911	324
Hydroburg	666,981	1,995	927
Watertown	44,822	59	6

B.3 Matching Weather Data to Irregular Meter Read Periods

Given the use of year-period fixed effects, the inclusion of weather covariates is needed to control for within read period variation in water use due to seasonality. This is especially necessary for bimonthly pilots. However, given that meter reads are taken at different times throughout the month and are of irregular lengths, matching water reads to monthly or even weekly weather would require many assumptions. We instead pull daily maximum temperature and daily precipitation from National Oceanic and Atmospheric Administration's National Climatic Data Center, specifically from the Global Historical Climatology Network-Daily (GHCN-D) database. We match households to the nearest weather station with available data. Then by looping through each irregular read period, we find the average maximum daily temperature and the average daily precipitation. We also estimate the average cooling degree days (CLDD; the accumulated degrees above 65 degrees Fahrenheit) per day using the daily maximum temperature. We use an average measure of CLDD since CLDD has a better predictive power than average temperature, and we also need to control for the variation in period length, making CLDD per meter read inappropriate. With the same argument we created the average number of rainy days (days with more than 0.01 inches of rain) per day. Exploring other options and data sources, we collected daily evapotranspiration (ET) at the zipcode level from the California Irrigation Management Information System (CIMIS) through their Spatial CIMIS system. ET is the combined loss of water from evaporation and transpiration (like evaporation but also the consumptive water use of plant tissues). For healthy plants including turf grass and gardens, this water needs to be

replaced each day. Figure B.3 provides the monthly accumulation of ET from daily averages across the zipcodes covered by the respective pilots. The variation in ET is driven more by solar radiation, relative humidity, and cloud cover, and less by average daily temperature (Hidalgo et al., 2005). For this reason, we suspect it is a better measure to use than average temperature. However, daily ET is missing for several full weeks and months in Watertown, the utility with the fewest households and observations. See Table B.7 for the percent of water reads with missing weather data. For this reason we use average CLDD and average days with precipitation in meter reading periods in our regressions to control for within read period variation in water use. See Figures B.4 and B.5 for pilot level time series of monthly averages of temperature and precipitation. Table B.8 reports pilot level monthly averages of all weather data collected.

Table B.7: Percent of Water Read Observations Missing Corresponding Weather Data

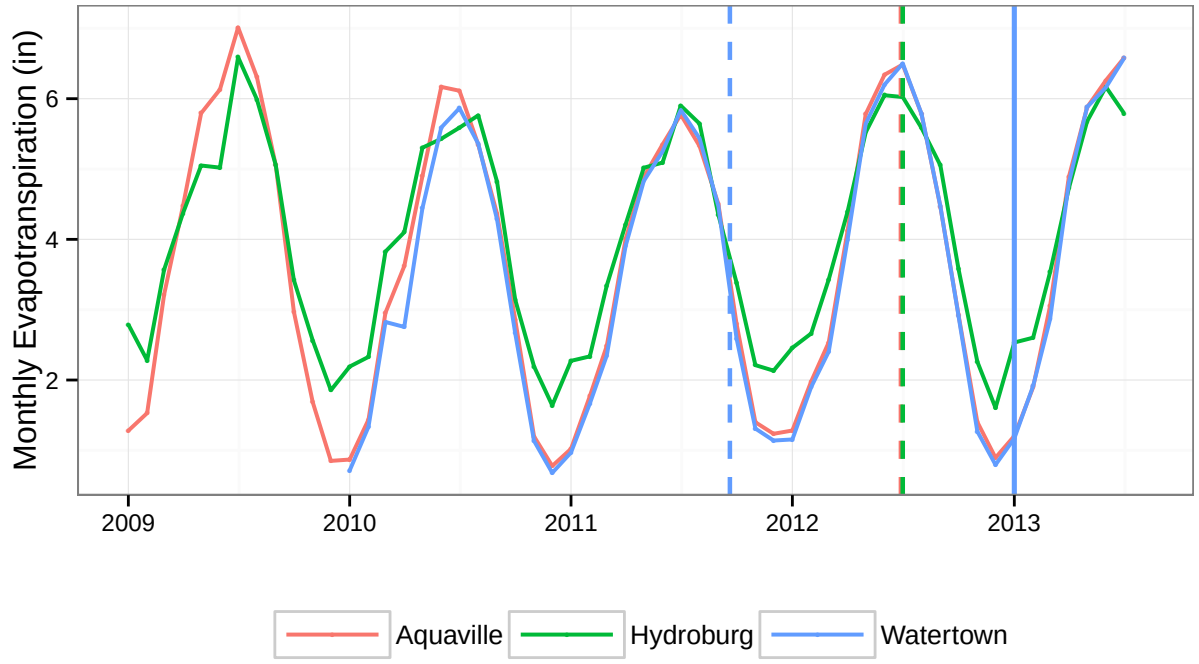
Pilot	Temperature	Evapotranspiration	Rain
Watertown	0.0	13.7	0.0
Aquaville	10.1	0.0	10.1
Hydroburg	0.1	0.1	0.1

Table B.8: Summary Statistics By Pilot: Baseline water use and monthly weather

Pilot	Water Use (GPD)	Temp (F)	CLDD	Rainy Days	ET (in)	Rain (in)
Watertown	202	70	138	5.5	3.5	2.2
Aquaville	282	67	49	7.0	3.7	1.7
Hydroburg	348	72	216	3.9	4.0	0.8

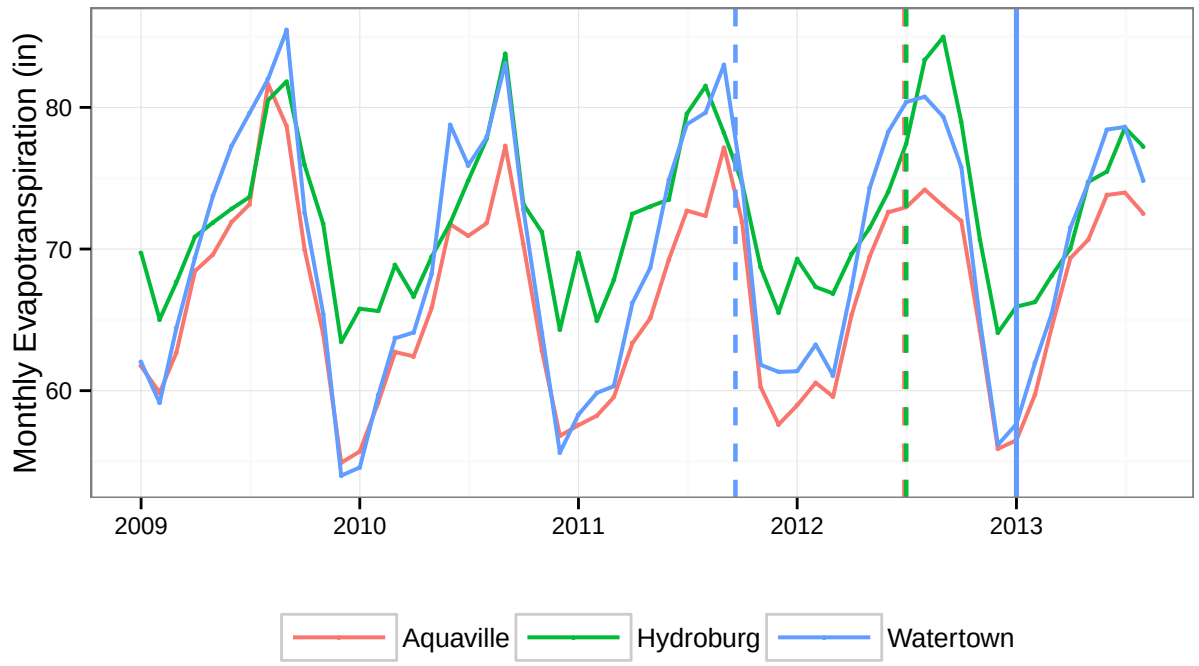
Notes: The water use reported is the mean gallons per day before for treatment across all households; Temperature is the average monthly maximum temperature in degrees Fahrenheit; CLDD stands for cooling degree days and is the average sum of degrees exceeding 65 degrees per month; Rainy days is the average number of days with more than 0.01 inches of rain per month; ET stands for evapotranspiration and is the combined loss of water from evaporation and transpiration (like evaporation but from plant tissues); Rain is the monthly average precipitation in inches.

Figure B.3: Evapotranspiration



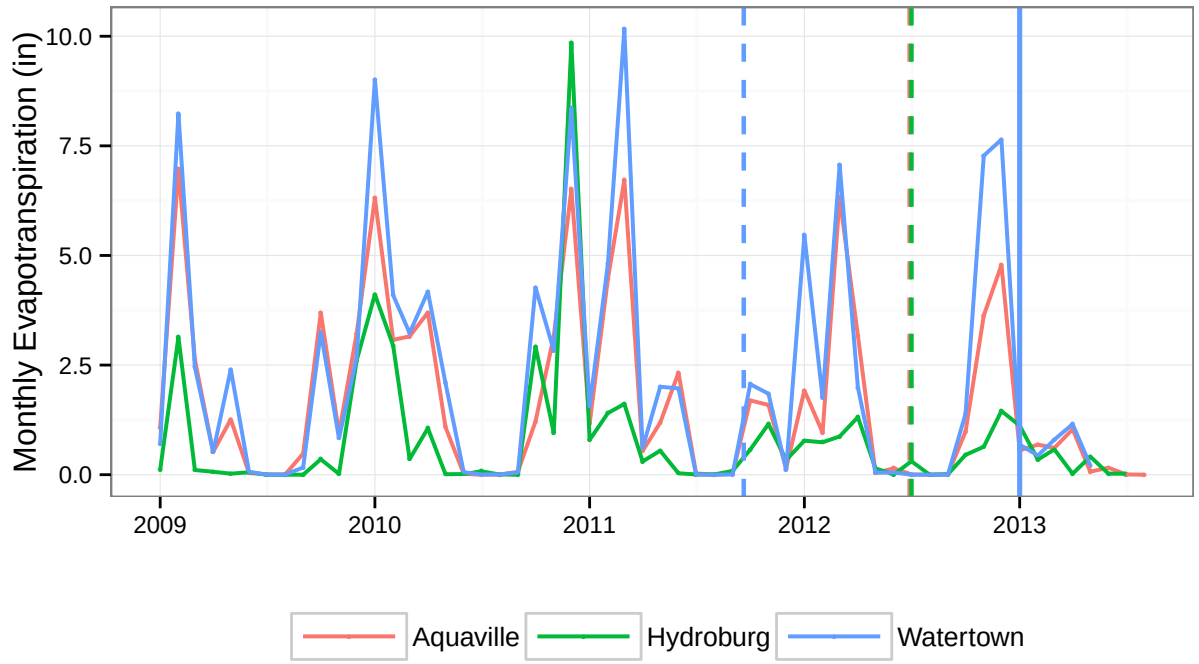
Notes: Evapotranspiration (ET) is the combined loss of water from evaporation and transpiration (like evaporation but from plant tissues). For healthy plants including turf grass and gardens, this water needs to be replaced each day. The graph provides the monthly accumulation of ET from daily averages across the zipcodes covered by the respective pilots. The variation in ET is driven more by solar radiation, relative humidity, and cloud cover, and less by average daily temperature **(SITE: Hidalgo et al. 2005)**. The vertical dashed lines indicate the start of the program for each pilot. The solid line indicates the end of the experiment for Watertown.

Figure B.4: Average Maximum Daily Temperature



Notes: The graph provides the average daily maximum temperature in the given months across multiple weather stations within the separate pilot locations. The vertical dashed lines indicate the start of the program for each pilot. The solid line indicates the end of the experiment for Watertown.

Figure B.5: Monthly Precipitation



Notes: The graph provides the monthly accumulation of precipitation from the average daily precipitation observed from multiple weather stations within the separate pilot locations. The vertical dashed lines indicate the start of the program for each pilot. The solid line indicates the end of the experiment for Watertown. In Dec 2010, Southern California experienced record storms causing flood and land slides.

B.4 Cleaning other outside data sources

We use block level median housing values from Zillow, an on-line real estate company with a database of millions of homes and sales transactions across the US.[2] This is meant to proxy for household income which is only available at the block-group level from the five-year American Community Survey. Also from Zillow's extensive database, we pull median year built and lot size at the census block level to replace missing values of these variables provided at the individual level from WaterSmart. With the exception of lot size in Hydroburg, a vast majority of the values used are at the individual level from WaterSmart. Table B.9 provides the percent of non-missing observations of these covariates that were provided by WaterSmart. We clean all home feature covariates by replacing values with missing if the value exceeds the median $\pm 10\times$the standard deviation within pilot.

Table B.9: Percent of Non-missing Home Feature Data Provided by WaterSmart

Pilot	Year Built	Lot Size
Aquaville	99.8	99.8
Hydroburg	100.0	0.1
Watertown	90.6	92.5

Notes: The remainder of the non-missing data are referred from block level medians using the Zillow datsbase.

The last external data source are voting records from the Statewide Database hosted by the University of California, Berkeley to create ideological indices capturing the environmental sympathies of individual block groups.[3] The Green Ideology Index (GII) is formed using six census block level measures from the 2008 and 2010 California elections:

- the percent who voted Democrat for U.S. Senator in 2008 and Governor in 2010
- the percent who voted yes on Proposition 7 (2008), which would have required Cali-

[2]See http://www.zillowblog.com/research/2012/01/21/zillow-home-value-index-methodology/ for information over the Zillow Home Value Index.

[3]The data are publicly available at http://statewidedatabase.org/.

fornia utilities to produce half their electricity from renewable resources by 2025

- the percent who voted yes on Proposition 10 (2008), which would have allocated $5 billion as cash incentives for high fuel economy and alternative fuel vehicles and R&D for and education on renewable energy and alternative fuel technologies

- the percent who voted yes on Proposition 21 (2010), which would have increased vehicle license fees in the state by $18 in order to raise $500 million a year dedicated to California State Parks

- the percent who voted no on Proposition 23 (2010), which would have suspended the state's 2006 Global Warming Solutions Act until the state's unemployment rate decreased to 5.5% for four consecutive quarters, a condition not met for decades.

For each measure, we assign a score near 100 to the top 0.1% of blocks with the greatest percentage of individuals voting along green-friendly lines and a score near zero to the least green-friendly census block. To ensure data quality, we only assign and use scores to blocks with more than thirty votes. Averaging these ordinal scores together yields the GII and maintains the 0 to 100 scale across all of California (the greenest areas are consistently green across all votes). To ensure data quality, we only maintain score averages (index values) where the census block has at least five of the six scores from the separate measures. We match individual households to the GII via their census block as a proxy for their own ideology under the assumption that like-minded individuals tend to cluster around green-related amenities that cater to their preferences.

Tables B.10 and B.11 provide pilot level summaries and data coverage for these descriptive covariates.

Table B.10: Summary Statistics By Pilot: Baseline water use and HH characteristics

Pilot	Water Use	Green Index	Income	Home Value	Year Built	Lot Size
Watertown	202	71	57,333	358,394	1987	5,227
Aquaville	282	62	96,696	670,242	1952	5,600
Hydroburg	348	34	120,943	815,231	1994	5,100

Notes: The water use reported is the mean gallons per day before for treatment across all households; Green Index is the median Green Index across pilot households, ranging from 0 to 100. The household index value is inferred from the Green Index of the census block. See the text for the creation of the block level index; Income is the median income across households in the pilot program. Individual household income is inferred from the average reported income in the census block; Home value is the median home value across households in the pilot program. Individual household home value is inferred from the median home value in the census block; Year build is the median year homes were built across pilot households. When not provided by WaterSmart, year built is inferred from the census block median; Lot size, in square feet, mirrors the logic of year built.

Table B.11: Percent of Households Missing Data

Pilot	Green Index	Income	Home Value	Year Built	Lot Size
Watertown	28.3	0.0	5.4	0.1	8.5
Aquaville	36.2	0.0	2.7	0.0	0.1
Hydroburg	22.3	0.1	12.1	0.1	13.9

Notes: Non-missing data for year built and lot size include those inferred from regional aggregates.

B.5 Water Use

B.5.1 Distribution and Outliers

There are many gallons per day observations that seem beyond reasonable. Table B.12 provides water use by pilot and by selected quantiles. We consider observations to be outliers, and run versions of our regression results excluding them, if the gallons per day exceed the $median + 8 \times interquartile\ range$ (within pilot and season). The assumption is that this procedure only drops water reads due to leaks or broken pipes and not the observations of heavy-handed water users. The outlier counts and water distributions by season are given in Figure B.6. Table B.13 provides seasonal average water use by pilot.

We also consider identifying outliers within households. In this case, there are too few

Figure B.6: Distributions of gallons per day by season

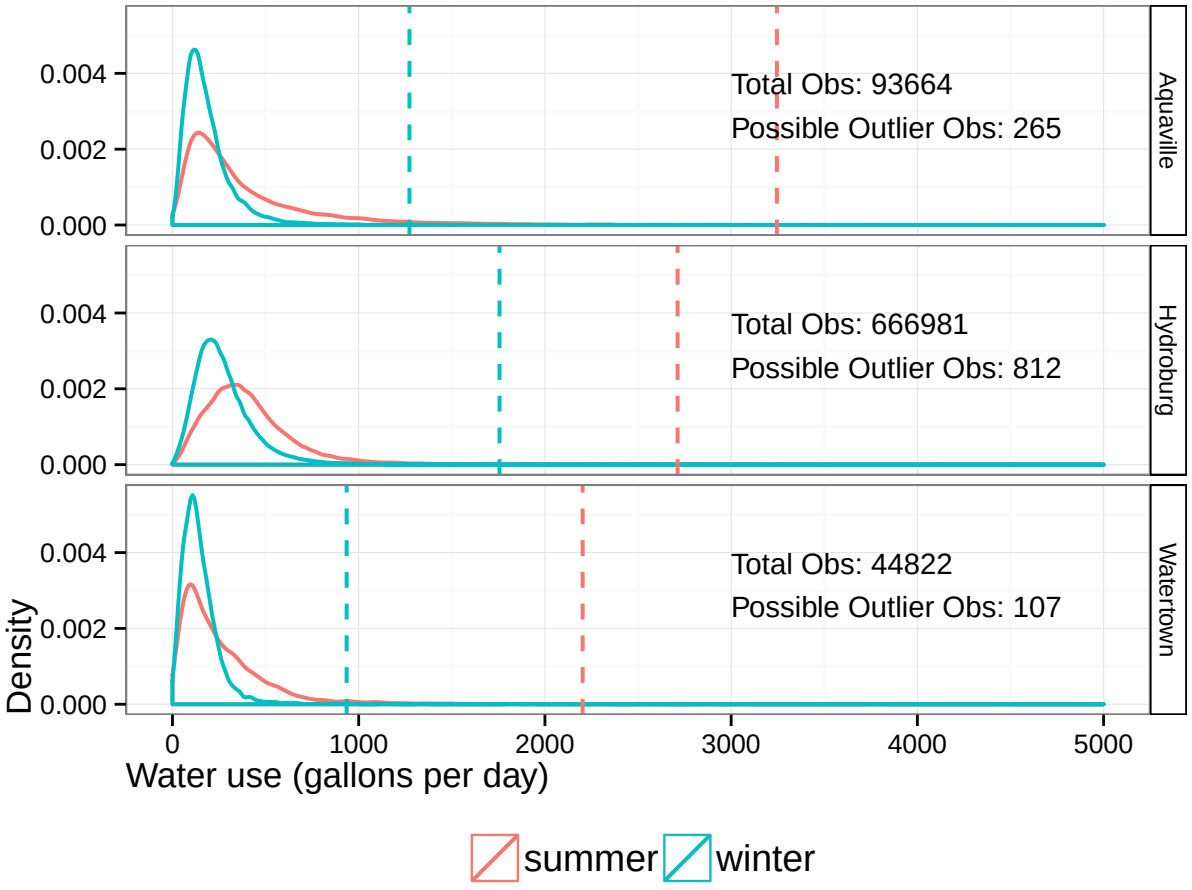


Table B.12: Water Use: High and low quantiles by pilot

	Watertown	Aquaville	Hydroburg
0th	0	5	19
1st	16	25	48
5th	33	51	97
25th	89	123	199
50th	150	196	299
75th	250	338	434
95th	531	811	748
99th	967	1521	1169
100th	4484	5800	7012

Notes: Water use is reported in gallons per day.

Table B.13: Median Water Use by Season

Pilot	Summer	Winter	Ratio
Aquaville	282.0	156.8	1.8
Hydroburg	374.0	249.3	1.5
Watertown	193.5	125.0	1.5

Notes: Water use is reported in gallons per day.

within HH and season observations as of now to determine without household outliers by season. Also, within HH distributions are wider (less n) so we also dial back the requirement. By this we mean that we consider daily water use to be too high or too low for the household if the water use exceeds the household median gallons per day $\pm 6\times$ the interquartile range of the household. Table B.14 and B.15 provide the number of observations defined as outliers for each of the rules described above as well as the number of observations identified under multiple rules.

B.5.2 Balance in Baseline Water Use Between Treatment and Control

In all three pilots, water accounts were randomly selected into the treatment group. A randomized field experiment facilitates estimation of the treatment effect by allowing us make the following assumptions. First, no selection bias exists: the receipt of a Home Water

Table B.14: Outliers in High Water Use (observation counts)

Pilot	Tot Obs	High 4 HH	High 4 Pilot	High 4 Pilot & High 4 HH
Aquaville	93,664	191	265	15
Hydroburg	666,981	821	812	73
Watertown	44,822	157	107	16

Table B.15: Outliers in Low Water Use (observation counts)

Pilot	Tot Obs	Below 20 GPD	Low 4 HH	Below 20 Gallons & Low 4 HH
Aquaville	93,664	217	18	6
Hydroburg	666,981	1	4	0
Watertown	44,822	1,007	13	9

Report must not depend on the change in water use before and after treatment. Households in neighborhoods selected to receive HWRs are no more likely than control households to invest in water saving infrastructure or engage in conservation behaviors in the absence of treatment. This could mean that relative to control households treatment households are equally susceptible to peer pressure, equally sympathetic to environmentalist causes, or equally lacking costly information on their costs of water use reduction. Randomization ensures the independence between treatment and any observed or unobserved covariates that may confound the estimation of treatment effects.

However, we may need to perform additional cleaning on the randomized pilots when analyzing heterogeneity in the response to treatment according to home values, environmental idealism, and pre-treatment water use. This is because randomization was not conducted within such subgroups, only across the entire population leaving room for imbalance within subgroups and potential spurious correlations (Ferraro, 2013). Balance refers to the degree in which the covariates' multidimensional distribution of the treatment group resembles the

multidimensional distributions of the control group. With more than three covariates, it is very difficult to inspect these multi-dimension distributions. Also, because balance is a property of the observed sample and not of of a hypothetical population there is no hypothesis test for balance. While we can never be sure we've achieved perfect balance, we can find evidence of imbalance by examining the first moments of individual covariates. Table B.16 shows that baseline water use is significantly larger in the control group than in the treatment by almost 14 gallons per day. Home values between treatment and control groups are also significantly different in Watertown, but due to the smaller sample size, Watertown is not a good candidate for matching procedures.

Table B.16: Differences in Means Between Control and Treatment Groups

(a) Variables used to explore heterogeneity

Pilot	Baseline Water	Green Index	Home Value	Income
Aquaville	-1.211 (0.861)	0.3 (0.48)	-12,107 (0.946)	-84 (0.77)
Hydroburg	13.89 (0)	2.7 (0.606)	109,701 (0.979)	3,337 (0.011)
Watertown	-6.541 (0.993)	0.5 (0.195)	-19,930 (0)	-530 (0.21)

(b) Structural Features of the House

Pilot	Sqft	Lot Size	Single Family Home	Year Built
Aquaville	-11 (0.753)	-178.4 (0.116)	0 (0.112)	1.7 (0.538)
Hydroburg	208 (0.292)	217 (0.738)	-0.015 (0)	8.5 (0.495)
Watertown	-32 (0.123)	-124.6 (0.854)	-0.029 (0.126)	1.2 (0.079)

Notes: Values reported are the control mean minus the treatment mean. P-values for the difference from zero are in parenthesis. While we don't use income explicitly to explore heterogeneity, the use of home value is used as a proxy because of its finer geographical granularity.

On Hydroburg, we employ a matching algorithm developed by Gary King[4] to produce a control sample as similar to treated households in pretreatment water use and select covariate space as possible to reduce or eliminate the bias from imbalance between treatment and control groups. This algorithm samples from the full set of control observations to minimize

[4]Documentation can be found at http://gking.harvard.edu/files/gking/files/cem.pdf.

the distance in covariate space between the sampled control set and the original treatment group. The specific covariates used for matching are the same observed variables that would facilitate the controlled prediction of the response variable, i.e. they directly effect the response variable as well as effecting treatment. In the case of this analysis which employs Difference-In-Difference (DID) regressions to estimate impacts on water, the response variable is ultimately water use after treatment. Important predictors then include pre-treatment water use (captured by household average baseline water use as well as seasonal averages in baseline water use), irrigable area, and year built (a proxy for the age and water-intensity of appliances). While other covariates may also be important, this set is more readily available for individuals in the Hydroburg pilot. Before running the algorithm, we drop control households without sufficient observations to overlap with the observations drawn from the treatment group.

QQplots help to illustrate the balance achieved by the matching algorithm and can be found in Figure B.7.

After applying this matching algorithm the difference between treatment and control group baseline water use remains, but the magnitude of this difference at the very least is reduced to only 0.151 gallons, see Table B.17.

Table B.17: Differences in Means Between Control and Treatment Groups for Hydroburg after Matching

(a) Variables used to explore heterogeneity

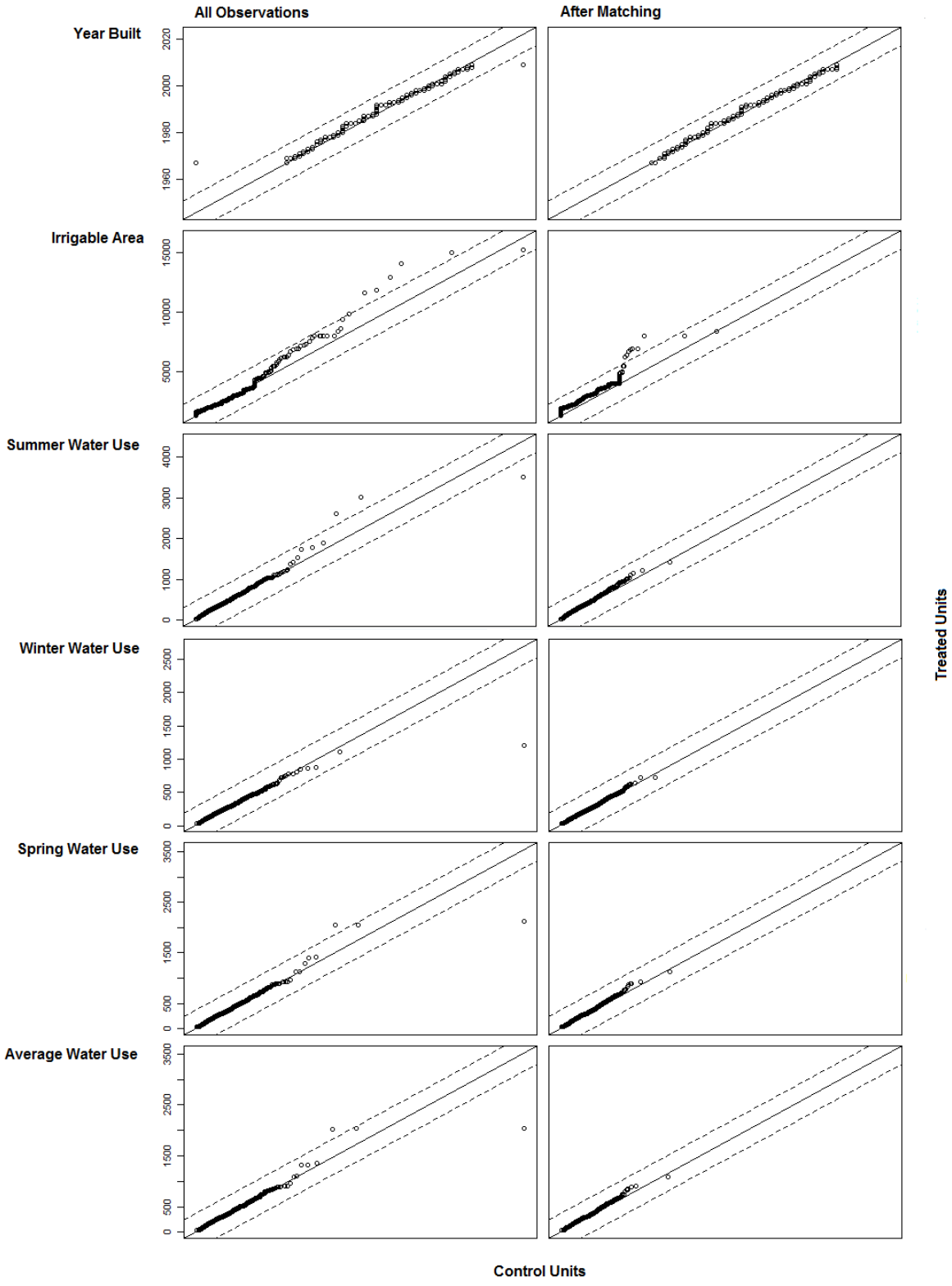
Pilot	Baseline Water	Green Index	Home Value	Income
Hydroburg	0.151 (0)	2.1 (0.319)	98,329 (0.651)	2,796 (0)

(b) Structural Features of the House

Pilot	Sqft	Lot Size	Single Family Home	Year Built
Hydroburg	177 (0.204)	75 (0.88)	-0.014 (0)	7.6 (0.843)

Notes: Values reported are the control mean minus the treatment mean. P-values for the difference from zero are in parenthesis. While we don't use income explicitly to explore heterogeneity, the use of home value is used as a proxy because of its finer geographical granularity.

Figure B.7: Quantile-quantile plots to explore effect of matching algorithm on balance

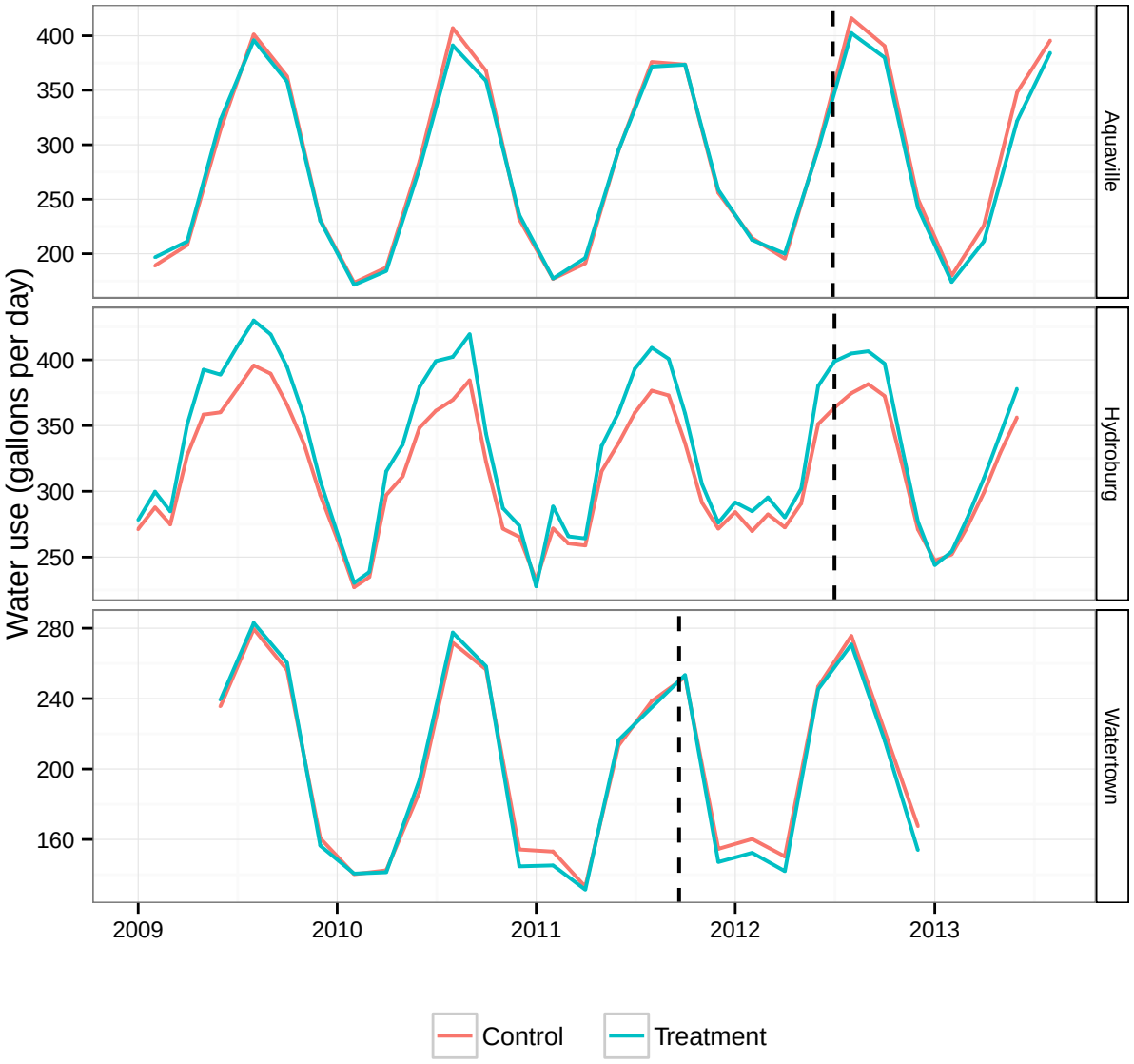


We also need to assume there are no treatment spill overs between treated and non-treated neighbors. The best we can do to avoid this in these random experiments is to drop the control households who have somehow found their way to the WaterSmart web portal feature without receiving their own HWR. This occurs for only a few households in Watertown. We know this because WaterSmart provides information on their site traffic.

B.5.3 Water Use Time Series

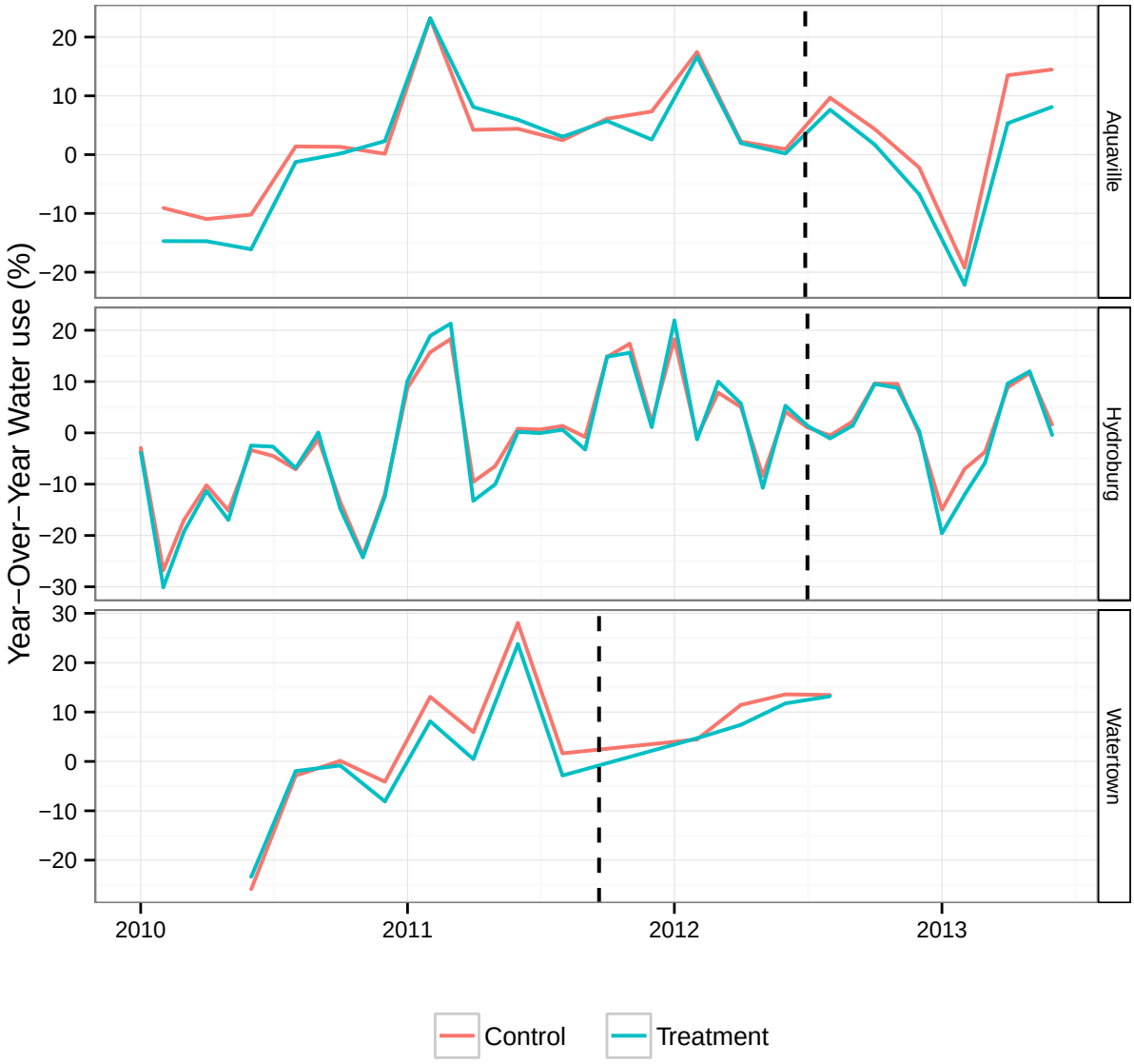
Figures B.8 provides time series of water use in gallons per day and Figure B.9 provides time series of year-over-year percentage changes in water use for each pilot.

Figure B.8: Water Use by Pilot



Notes: The graph contains the mean water used per day in gallons for both the treatment and control groups over time. The vertical dashed line indicates the start of the program for each pilot.

Figure B.9: Year-Over-Year Change in Water Use by Pilot



Notes: The graph contains the year-over-year change in water use in percentage terms for both the treatment and control groups over time. The vertical dashed line indicates the start of the program for each pilot.

B.6 Conditional Average Treatment Effects

This section presents the full output for regressions that interact the treatment effect with covariates to identify heterogeneity in conditional average treatment effects. Tables B.18, B.19, and B.20 contain the regression that produce Tables 2.8, 2.9, and 2.10 respectively that display linear combinations for the treatment effect and the treatment effect interaction.

Table B.18: Heterogeneity: Baseline Water Use

	(1) Watertown	(2) Aquaville	(3) Hydroburg
Treatment Effect	-0.0537**	-0.0732***	-0.0509***
	(0.0237)	(0.0195)	(0.0138)
TE*Quintile 1	0.0487	0.0945***	0.0716***
	(0.0338)	(0.0274)	(0.0234)
TE*Quintile 2	0.0433	0.0387	0.0947***
	(0.0291)	(0.0269)	(0.0206)
TE*Quintile 4	-0.0316	0.0371	0.0060
	(0.0274)	(0.0247)	(0.0187)
TE*Quintile 5	-0.0561**	0.0366	0.0094
	(0.0262)	(0.0262)	(0.0200)
Rainy Days	22.1677	-0.3019***	-8.5055***
	(19.9632)	(0.0367)	(0.3943)
Cooling Degree Days	0.2068	0.0067***	0.0013
	(0.6671)	(0.0009)	(0.0047)
Household FEs	Yes	Yes	Yes
Year-Period FEs	Yes	Yes	Yes
Adjusted R^2	0.162	0.251	0.190
Households	2,232	3,091	11,696
Observations	43,541	83,340	664,075

Notes: The dependent variable is the natural logarithm of water consumption in gallons per day. Robust standard errors clustered at the household level are reported in parentheses. Interaction variables are the treatment effect multiplied by quintiles of water consumption averaged over the entire pre-treatment period. The 3rd quintile is omitted, so the base effect can be interpreted as the effect at the median of baseline consumption. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.19: Heterogeneity: Ideology

	(1) Watertown	(2) Aquaville	(3) Hydroburg
Treatment Effect	-0.0533*	-0.0458**	-0.0148
	(0.0281)	(0.0231)	(0.0156)
TE*Quintile 1	-0.0149	0.0306	-0.0086
	(0.0330)	(0.0280)	(0.0225)
TE*Quintile 2	0.0262	0.0654**	-0.0531**
	(0.0338)	(0.0284)	(0.0222)
TE*Quintile 4	-0.0101	0.0200	-0.0037
	(0.0383)	(0.0291)	(0.0228)
TE*Quintile 5	-0.0397	-0.0058	0.0236
	(0.0377)	(0.0344)	(0.0258)
Rainy Days	25.7393	-0.3302***	-8.2944***
	(24.1518)	(0.0446)	(0.4408)
Cooling Degree Days	-0.1411	0.0075***	-0.0030
	(0.7071)	(0.0011)	(0.0049)
Household FEs	Yes	Yes	Yes
Year-Period FEs	Yes	Yes	Yes
Adjusted R^2	0.162	0.290	0.203
Households	1,841	1,958	9,024
Observations	32,110	54,869	516,534

Notes: The dependent variable is the natural logarithm of water consumption in gallons per day. Robust standard errors clustered at the household level are reported in parentheses. Interaction variables are the treatment effect multiplied by quintiles of the green ideology index measured at the census block level. The 3rd quintile is omitted, so the base effect can be interpreted as the effect at the median ideology. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.20: Heterogeneity: Housing Values

	(1) Watertown	(2) Aquaville	(3) Hydroburg
Treatment Effect	-0.0593***	-0.0372*	-0.0251
	(0.0216)	(0.0216)	(0.0155)
TE*Quintile 1	-0.0069	-0.0333	0.0075
	(0.0308)	(0.0294)	(0.0229)
TE*Quintile 2	0.0422	-0.0330	-0.0026
	(0.0304)	(0.0310)	(0.0234)
TE*Quintile 4	-0.0028	0.0264	-0.0044
	(0.0273)	(0.0259)	(0.0204)
TE*Quintile 5	-0.0010	0.0568**	0.0189
	(0.0262)	(0.0248)	(0.0217)
Rainy Days	25.2592	-0.3032***	-8.9716***
	(20.2693)	(0.0374)	(0.4406)
Cooling Degree Days	0.4395	0.0068***	0.0168***
	(0.6802)	(0.0009)	(0.0050)
Household FEs	Yes	Yes	Yes
Year-Period FEs	Yes	Yes	Yes
Adjusted R^2	0.162	0.255	0.210
Households	2,417	3,011	10,170
Observations	42,368	81,034	584,217

Notes: The dependent variable is the natural logarithm of water consumption in gallons per day. Robust standard errors clustered at the household level are reported in parentheses. Interaction variables are the treatment effect multiplied by quintiles of housing values averaged at the census block level. The 3rd quintile is omitted, so the base effect can be interpreted as the effect at the median average housing value. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$