THE UNIVERSITY OF CHICAGO

BELIEFS AND PERCEPTIONS IN ENERGY AND ENVIRONMENTAL ECONOMICS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
EUNKI MIN

CHICAGO, ILLINOIS
AUGUST 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am indebted to my thesis committee, Michael Greenstone, Leonardo Bursztyn, and Fiona Burlig, for their invaluable guidance and support. I am truly privileged to have them as my mentors and my teachers.

I was also fortunate to have studied at the University of Chicago. My colleagues played a crucial role in developing and enriching this dissertation. I am grateful to Tae Ho Kim, for early development of the first chapter, Hee Kwon Seo, for discussions on the second chapter, and James Traina, for priceless insights throughout the years. I appreciate warm comments from Eric English, Antonio Gabriel, Andrew Gianou, Kai-Wei Hsu, Amir Jina, Karthik Nagarajan, Ishan Nath, Hieu Nguyen, Hyunmin Park, Luis Simon, and Yuyao Wang.

I received enormous support from various institutions and intellectual communities. I thank NORC at the University of Chicago for the General Social Survey data, Energy Policy Institute at the University of Chicago (EPIC) for cultivating an excellent research environment, University of Michigan and Columbia University for the opportunity to share my research, and UC Berkeley for hosting Energy Camp for young energy economists.

Another gift for me is that I met amazing people in graduate school and became friends with them. I cannot imagine how I would have navigated through the core exam and beyond without my first-year study buddies, Anna-Maria, Dan, David, James, Kai Hao, Kai-Wei, Yuyao, and Ziho. Together we spent many days and nights solving problems, discussing economics, or having fun.

My parents have been huge supporters throughout my life, and they have always treated my sister and me with respect and care. The older I grow, the more I realize how challenging it is to be a good parent like them. Mom and dad, thank you!

The best thing in my life happened in Chicago, where I met my wife and my daughter was born. They were always there for me even during the toughest times, and they taught me what true love and happiness are. Seoyeon and June, I am excited to begin the new chapter of our life with you, and I cannot thank you enough for being my family.

ABSTRACT

Policies are potentially suboptimal when political opinions are not based entirely on scientific fact. Chapter 1 demonstrates that local industry composition has a substantial causal effect on environmental policy opinions and election outcomes. I show that county-level employment in the mining and extraction sector (NAICS 21) reduces support for environmental policies. I use the fracking revolution as a quasi-experiment that provides a clear, exogenous shock to mining employment. At the individual level, I find that being in or having a family member in the mining sector shapes environmental beliefs. Economic conditions further influence election outcomes; employment in mining increases vote shares for Republican candidates.

In Chapter 2, I study consumers' behaviors when they experience nonlinear prices. Notches, or discontinuous jumps in total prices, are common in tax and regulation, generating interesting theoretical and behavioral responses. Nonlinear budget sets induced by notches and kinks, or discontinuity in slopes of total prices, allow researchers to estimate structural parameters. Although most studies on notch points and nonlinear prices focus on the labor market and income contexts, I exploit a unique price scheme and a major reform in the Korean residential electricity market to study notches in consumption goods markets. Extant studies with milder forms of nonlinear prices (i.e., kinks) find that consumers do not internalize marginal prices. Although my context is more extreme in the sense that both total and marginal prices jump, and marginal prices vary up to 11.7 times, I find that consumers still do not perceive jumps in total payment.

CHAPTER 1

DO ECONOMIC CONDITIONS SHAPE ENVIRONMENTAL BELIEFS?

1.1 Introduction

Environmental beliefs and policies are not based entirely on scientific evidence and are affected by voters' interests. Figure 1.1 illustrates that regional variation in the environmental opinions is substantial, whereas the environmental opinions would have been homogeneous across regions if these opinions were based merely on scientific facts. Moreover, only 70% of Americans acknowledged global warming as of 2018, but climate scientists reached near consensus on the presence and consequences of climate change.[1] Since policies based on non-scientific beliefs might result in suboptimal outcomes, social scientists have studied obstacles in science communication.

Whereas primary explanations have suggested that party identity[2] or experiences[3] influence environmental beliefs,[4] these factors do not fully explain how environmental beliefs are shaped. For instance, Figures 1.2 and 1.3 show that region-specific factors other than party ideology affect environmental opinions. Although party ideology does explain average differences in attitudes toward climate change, geographic patterns of environmental beliefs in Figure 1.1 persist even after controlling for party ideology.

Indeed, a growing literature suggests a role played by economic conditions. Kahn and Kotchen (2011) finds that state-level unemployment correlates negatively with interest in

1. Meta-studies on the consensus on climate change among scientists include Cook et al. (2013), Cook et al. (2016), and Rosenberg et al. (2010).

2. Druckman, Peterson, and Slothuus (2013), Bolsen, Druckman, and Cook (2014), Leeper and Slothuus (2014), and Bolsen and Druckman (2018) illustrate that party identity has influences even on solely scientific opinions that have less to do with politics per se.

3. Howe et al. (2014), Demski et al. (2017), and Borick and Rabe (2017) explain that experiences or exposure to weather events caused by climate change affect environmental opinions

4. Other explanations include Gifford (2011) and Deryugina and Shurchkov (2016).

climate change. Mildenberger and Leiserowitz (2017) argues that shifts to political cues are primary determinants of environmental opinions, and business cycles, media coverage, and weather experiences have minimal influences on policy opinions. Olson-Hazboun, Howe, and Leiserowitz (2018) suggests that industry composition and activity in industries related to the environment are more pertinent than are measures of overall business cycles; mining activity relates to reduced support on renewable energy policies, which accords with extant findings that the reliance of local economies on specific industries shapes residents' political opinions[5]. Fedaseyeu, Gilje, and Strahan (2015) further demonstrate that conservative policies are supported more in areas with hydraulic fracturing activities. Thus, a correlation between industry composition and environmental beliefs is evident.

Despite consistent patterns in empirical findings, evidence of a causal link between industry composition and environmental opinions and policies is still limited. A primary challenge when estimating a causal relationship is that local environmental beliefs might affect mining firms' decisions to enter a region. A county favorable to mining activities attracts more mining companies and results in greater dependence on mining. This reverse causality has been overlooked due to a lack of clear exogenous shocks in industry compositions. One exception that isolates a causal link between economic conditions and policy opinions is Brunner, Ross, and Washington (2011), where the authors find that industry composition has a causal influence on support for redistributive policies, using exogenous shifts to labor demand. Their instrument is the Predicted Employment Index (PEI),[6] which predicts local employment using initial local industry composition and national industry employment growth rates. However, Goldsmith-Pinkham, Sorkin, and Swift (2018) warns that the PEI is equivalent to instrumenting for local industry shares, and it thus might be an inadequate instrument if the dependent variable, policy opinions in this case, influences local industry

5. Mukherjee and Rahman (2016), Boudet et al. (2016), Cragg et al. (2013), and Tvinnereim and Ivarsflaten (2016) explain that dependence on mining and drilling relates to environmental policy attitudes.

6. This index was developed by Bartik (1991) and popularized after Blanchard et al. (1992), and the main assumption that this index represents a valid instrument is that trends in national employment levels do not correlate with local responses to labor supply.

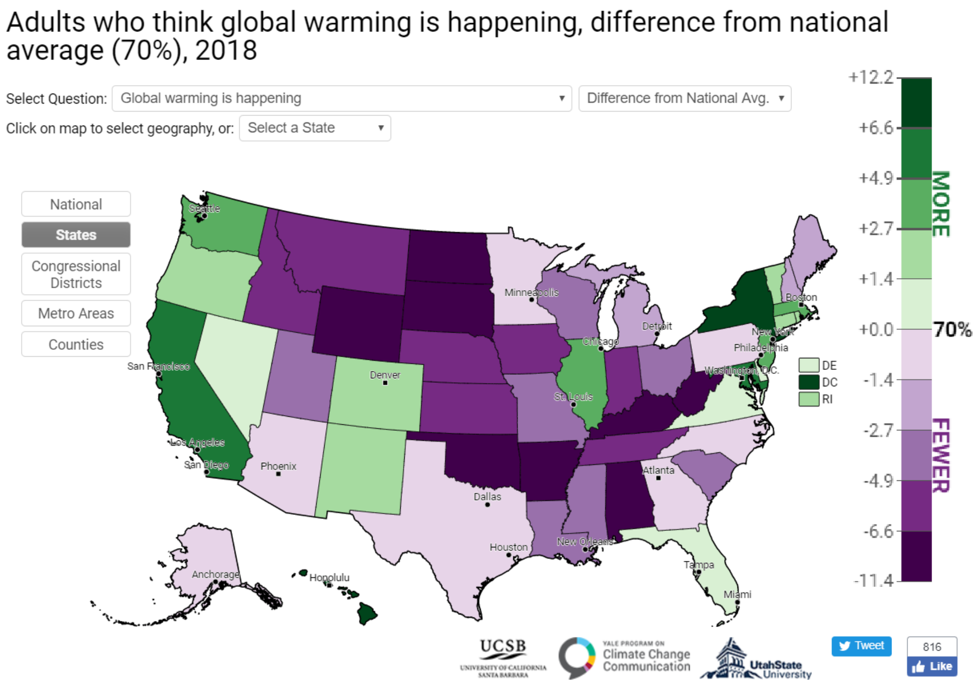


Figure 1.1: Opinions about climate change on state level, 2018

Note: This figure plots the regional variation in the opinions about climate change. The portion of American adults that believe climate change is happening in year 2018 is depicted. States in green color indicate that the portion of residents that believe global warming is happening is higher than national average, where states in purple believe less in climate change. The national average is 70% as of 2018.
Data Source: Howe et al. (2015)

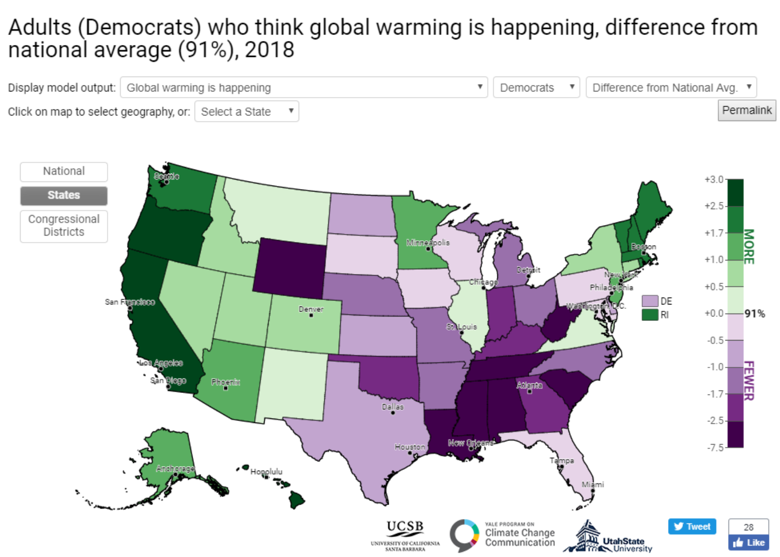


Figure 1.2: Democrats' opinions about climate change, 2018

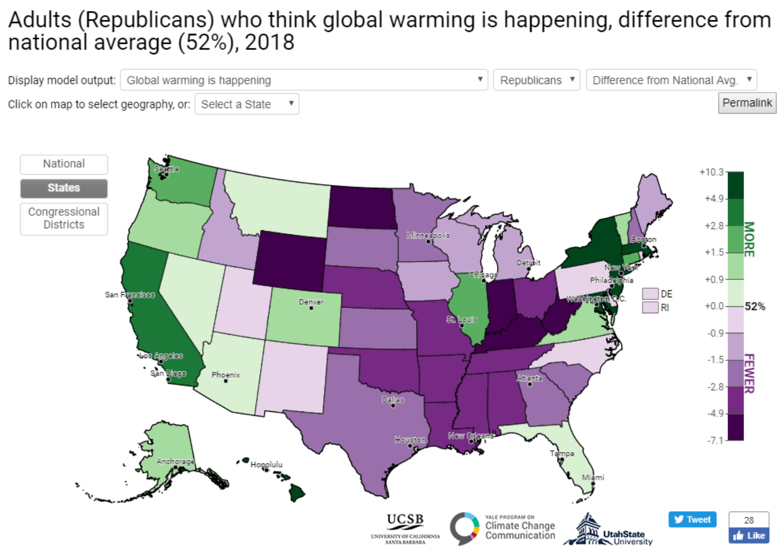


Figure 1.3: Republicans' opinions about climate change, 2018

Note: Figures 1.2 and 1.3 plot the respective portions of Democrats and Republicans that believe climate change is happening in year 2018. States in green color indicate that the portion of residents that believe global warming is happening is higher than national average, where states in purple believe less in climate change.
Data Source: Howe et al. (2015)

composition. I solve this identification problem using an alternative instrument based only on geological formations and technological shocks.

This paper assesses the mining industry's influence on environmental opinions and related policies. The fracking revolution, or the shale boom, provides a clear and exogenous shock to mining employment. I use the presence of shale formations to capture drilling properties of each shale reservoir and technological shocks specific to each shale formation. One assumption for exogeneity of the instrument is that environmental beliefs do not cause geological properties and changes to the profitability of shale formations. Use of shale geography as a quasi-experiment accords with Feyrer, Mansur, and Sacerdote (2017), Kearney and Wilson (2018), Bartik et al. (2019). Such research points out the economic significance of the fracking boom, suggesting the relevance of this instrument in this context. I find that the first stage is strong in the dataset; the fracking boom allows causal inferences regarding industry composition and policy opinions.

Using a special contract with one of the largest independent social research organizations in the United States, I construct a dataset of individual policy opinions and occupations. The dataset contains geographic IDs (i.e., Census tract, county, and state), environmental beliefs, previous votes, and occupations of individuals and family members. To measure the local industry composition of each individual, I link geographic locations with local economic characteristics that are published by the U.S. Census Bureau (Census Bureau 2019a,b). I map shale geography in the dataset using shale play shapefiles provided by the Energy Information Administration (EIA 2016). The combination of individual-level data with geographic ID matched with local economies and shale formulations allows me to examine how economic conditions affect opinions and preferences disaggregately.

Using shale geography, I find that the mining industry causally influences environmental opinions and election outcomes. If mining employs an additional 1 percentage point of total employment in a county, an average resident's skepticism about climate scientists increases by 4.2 percentage points and support for environmental protection decreases by 3.9 percentage points. Such employment also led to a 3.6 percentage point greater vote share for the Republican candidate during the 2008 presidential election. Since presidential elections can be close, it is possible that the influence of employment in each industry acts as a tie-breaker regarding votes. Being a miner has an even more drastic influence; miners are 55.1 percentage points less prone to support environmental protection.

By providing empirical evidence for a causal relationship, this paper contributes to literature on economic conditions and policy. I reaffirm a similar correlational pattern that the literature discusses. This paper serves as another example of studies that use the fracking boom as a quasi-experiment on local economies, and it builds on others that assess local responses to economic shocks (Allcott and Keniston 2017; Bartik 1991; Blanchard et al. 1992). This study also contributes to literature on climate denial[7].

7. See footnotes 2, 3, and 4 for studies on disagreements between the public and environmental scientists regarding climate change.

1.2 Fracking Boom as an Exogenous Shock to Local Economy

The hypothesis tested in this paper suggests that local industry composition influences residents' attitudes toward environmental policy. To assess the influence of local economies on environmental beliefs, ruling out reverse causality using an exogenous shock on economic conditions is required. Although there are many ways to study the effects of economies on policy opinions, I focus on the mining industry since a clear exogenous shock in this sector makes it feasible to estimate causal influences.

1.2.1 Problems in Identification: Endogeneity in Mining Decisions

Among a variety of measurements of a region's dependence on the mining industry, I use employment in the mining industry relative to total employment. Mining employment in a region discourages residents from discussing environmental issues since doing so might be harmful to employers and risks an individual's employment. Consequently, there are fewer opportunities to learn about environmental issues, and residents are uninterested about environmentalism.

When estimating the influence of mining employment on environmental beliefs, two sources of endogeneity are reverse causality and confounding factors. Employment in the mining industry is potentially a function of environmental policies. Potential confounders might influence both environmental beliefs and industry composition. By overlooking this aspect, results would be correlational and implications misleading.

Environmental beliefs affect a mining firm's decision to enter and extract resources in a region, leading to potential reverse causality. Mining employment is a function of both the geological presence of resources and mining firms' decisions to extract those resources. Environmental beliefs, or any economic variable generally, cannot alter the availability in a region. A region less interested in environmental protection would thus disproportionately attract more mining firms. Mining employment and environmental beliefs can also be caused by confounding factors, and other confounders, including culture and history related to mining, are unobservable and cannot be controlled directly. Due to the endogeneity issues described above, correlational patterns in data and extant research do not necessarily imply causal effects.

1.3 Fracking Boom as a Quasi-Experiment

Recent innovations in hydraulic fracturing, or fracking, and horizontal drilling technology led to increases in shale gas and oil production. The fracking boom drastically reshaped local economies; regions with shale formulations experienced increased employment and income, especially in the mining and drilling sector. I exploit the fracking boom as a quasi-experiment and discuss the relevance and exogeneity of the resulting instrument.

Shale gas and oil are unconventional sources of fossil fuels that are trapped in shale formations.[8] Contrary to conventional gas or oil wells that are trapped in rock formations in which fluids can pass freely (or permeable) and form larger pools, shale gas does not flow freely and remains in multiple smaller pockets.[9] Conventional gas and oil flow up through the ground by themselves or can be pumped inexpensively, but shale gas had not been extracted on a commercial scale until cost-effective ways to extract resources from unconventional wells were developed and became widespread. Figure 1.4 shows the difference between conventional and unconventional oil wells.

Beginning in early 2000, enhancements to hydraulic fracturing and horizontal drilling technologies started the fracking boom—the rapid increase in commercial extraction of shale gas and oil. Hydraulic fracturing, or fracking, is a technology that makes shale layers permeable so that shale gas and oil become extractable. During fracking, shale layers are perforated to increase permeability, and pressurized fluid is injected to release fossil fuels trapped in the

8. Energy Explained, (EIA 2019b): https://www.eia.gov/energyexplained/natural-gas/where-our-natural-gas-comes-from.php

9. Shale layers consist of particles finer than a few microns (μm), preventing liquids from moving inside.

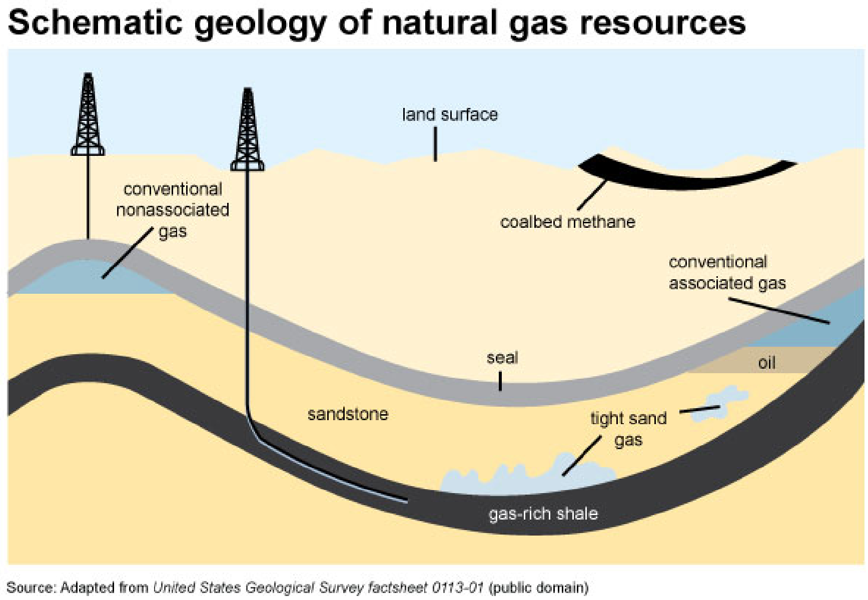


Figure 1.4: Conventional and Unconventional Oil Wells

Note: This Figure depicts typical structure of conventional and unconventional oil wells. Unconventional oil extraction involves in horizontal drilling and hydraulic fracturing procedures, as depicted in the picture.
Source: EIA (public domain)
https://www.eia.gov/energyexplained/natural-gas/where-our-natural-gas-comes-from.php

sediment. Since shale gas is often trapped in a horizontal layer, vertical wellbores turn and keep drilling horizontally to maximize contact with the surface area of the shale layer. With advances to fracking and horizontal drilling, extraction of previously inaccessible shale gas and oil became widespread. The increase to natural gas production since 2005, due largely to shale gas, led to energy independence in the United States. As of 2018, 69% of total U.S. natural gas production comes from shale gas. Shown in Figure 1.5, the rapid increase in shale gas is projected to grow continuously in the future.

Commercial production of shale gas had significant influences on local economies, especially regarding mining employment. Kearney and Wilson (2018) provide first-hand evidence that the jobs-to-population ratio increases in fracked areas. New shale oil and gas production increases employment in mining and drilling industries, according to Feyrer, Mansur, and Sacerdote (2017), an influence that persists; two-thirds of the increase persists two years after new shale gas and oil production. In a comprehensive analysis of the influence of

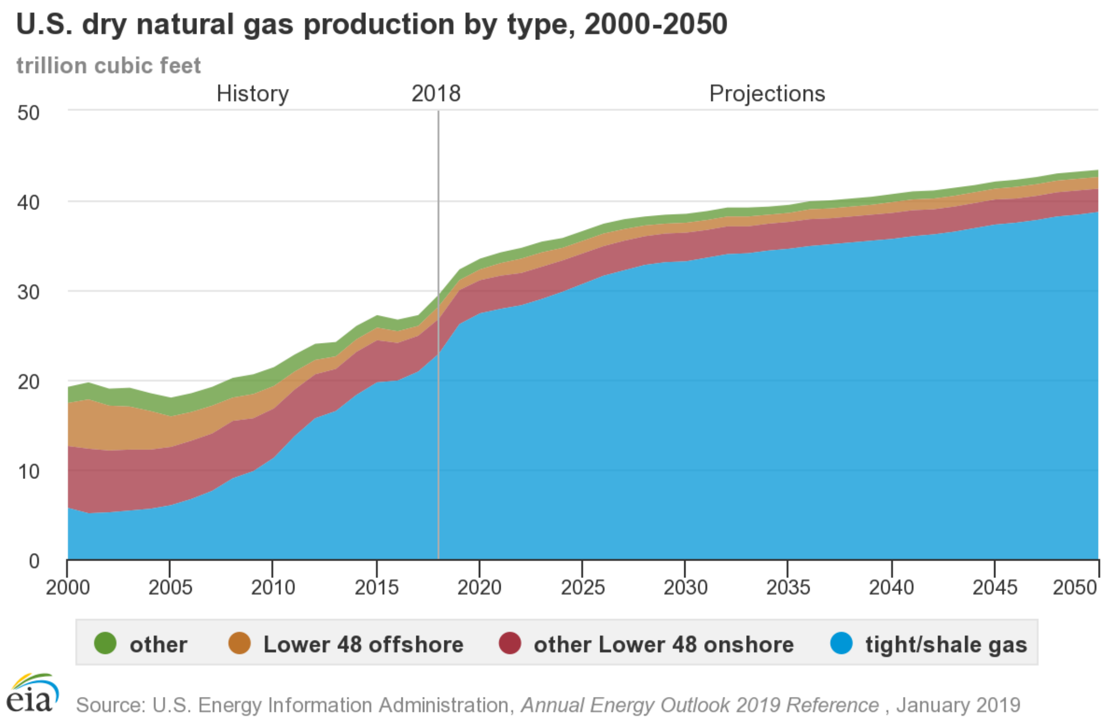


Figure 1.5: Continuous Increase in Shale Gas Production

Note: This figure plots the volume (in trillion cubic feet) of annual U.S. dry natural gas production by the type of wells. Tight/shale gas, in blue color, represents natural gas produced in unconventional wells. Shale gas production has drastically increased starting in early 2000's and surpassed the conventional gas production in volume.
Source: EIA (2019a)
`https://www.eia.gov/energyexplained/natural-gas/where-our-natural-gas-comes-from.php`

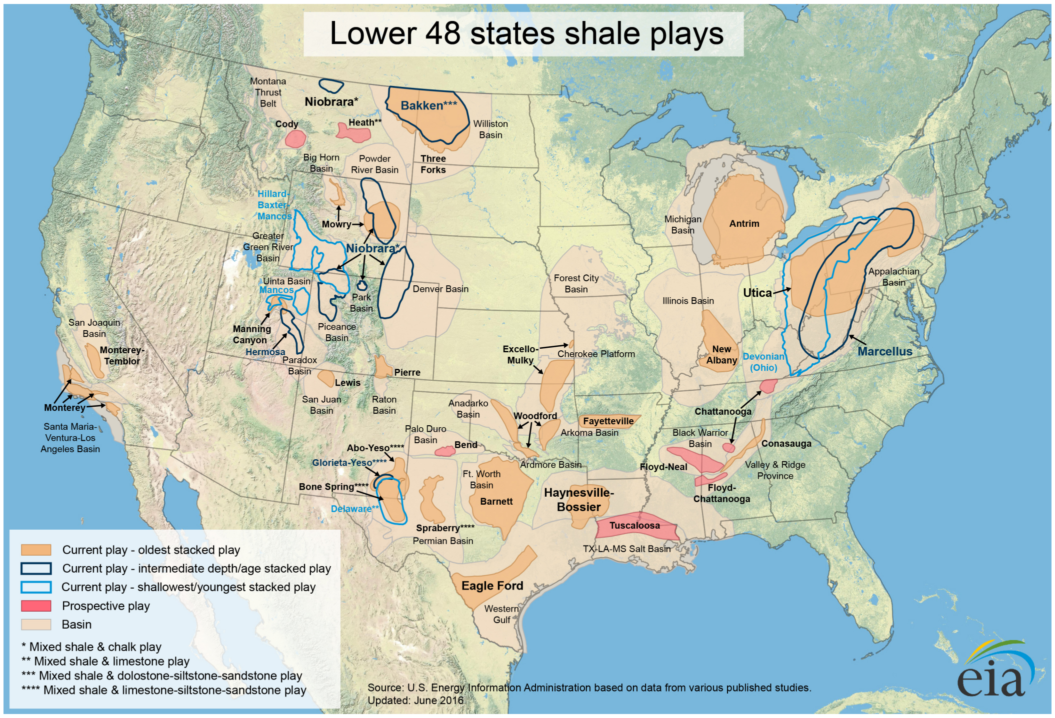


Figure 1.6: Shale Gas and Oil Plays in the Lower 48 States

Note: This figure plots shale gas and oil reservoirs in the lower 48 states.
Source: EIA (2016)
`https://www.eia.gov/energyexplained/natural-gas/where-our-natural-gas-comes-from.php`

fracking on local economies and welfare, Bartik et al. (2019) finds that mining employment increases over 40 percent, and fracking production increases net welfare even when decreased amenities are considered. The literature consistently finds that fracking changes economies and disproportionately affects the mining industry. Figure 1.6 illustrates regions potentially affected by the shale boom.

I reaffirm that the fracking boom has substantially affected local economy and increased mining employment. Figure 1.7 provides the mining employment variable, defined as the number of employees in mining industry relative to the total number of employees in each county, in counties on shale plays (or fracked counties) and the rest of the counties. In Figure 1.8, I further provide the estimated differences in mining employment between fracked and non-fracked counties. This figure illustrates that counties on shale areas faced significant increase in mining employment as fracking starts to dramatically increase in year 2004.

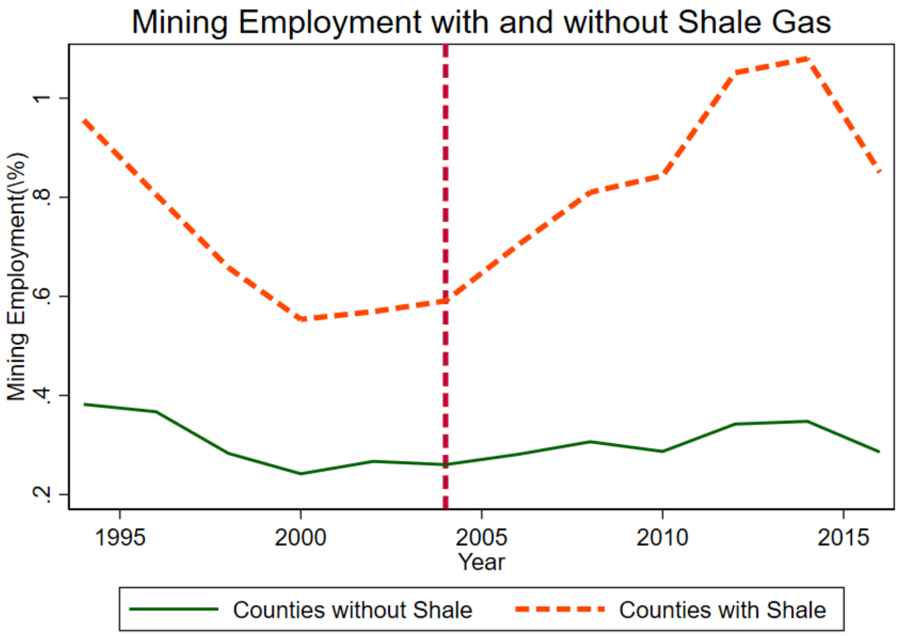


Figure 1.7: Mining Employment in Fracked and Other Counties

Note: This figure plots trends in mining employment in counties without and with shale resources. I use shapefiles for US shale regions provided by EIA (2016) to determine if a county has shale resources. The green solid line represents the employment in mining industry relative to the total employment in average counties that are outside shale gas regions. The orange dashed line represents the mining employment in counties with shale resources. Year 2004, marked with red vertical dashed line, indicates the year hydraulic fracturing began in commercial scale.
Source: Shapefiles of shale plays come from EIA (2016). Mining employment is computed using County Business Patterns, U.S. Census Bureau

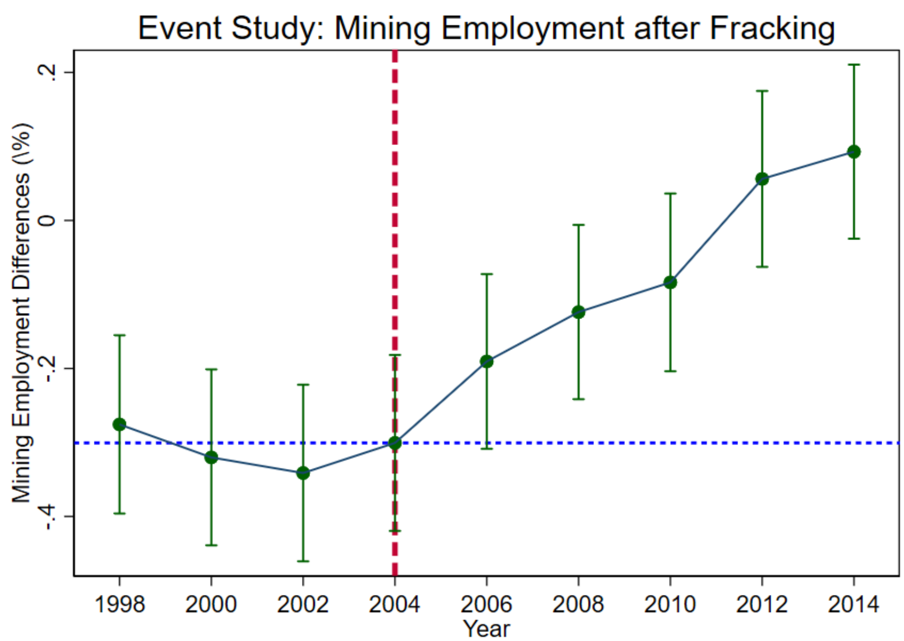


Figure 1.8: Event Study: Differences in Mining Employment between Counties with and without Fracking

Note: This figure plots the estimated differences in mining employment between counties with and without shale resources. I use shapefiles for US shale regions provided by EIA (2016) to determine if a county has shale resources. I control for year fixed effects and the time-invariant fixed effects of being on shale plays, then estimate the differences in mining employment between counties with and without shale for each year. The reported coefficients are $\hat{\delta_j}$ in the following equation:

$$y_{it} = \text{Shale Counties}_i + \text{year FE}_t + \sum_{j \neq 2004} \delta_j (\text{Shale Counties}_i \cdot I(t = j)) + \epsilon_t$$

The point estimates for the differences reported as dots, and 95 percent confidence intervals are shown as green vertical spikes. Year 2004, marked with red vertical dashed line, indicates the year hydraulic fracturing began in commercial scale. Whereas mining employment did not significantly differ before 2004, counties on shale plays face significant increase in mining employment after the fracking boom. I use number of total employees in each county as weights in the above regression.
Source: Shapefiles of shale plays come from EIA (2016). Mining employment is computed using County Business Patterns, U.S. Census Bureau.

While the differences between the mining employment is relatively constant before year 2004, mining employment significantly increases in the fracked counties after the fracking boom.

Shale formations can have disparate geological characteristics. Sub-formations of shale that share geological and drilling properties are defined as shale plays, which are located in larger-scale shale formulations called shale basins. A well's productivity is determined by depth, thickness, the gas-to-oil ratio, porosity, percentage of organic content, and other geological characteristics (EIA 2011).

I exploit the availability of shale plays, which are plausibly exogenous, and variations in geological characteristics and productivity of each shale play of which counties have less control. Due to heterogeneous geological characteristics, some shale plays become recoverable only when adequate extraction technologies are used and crude oil prices are sufficiently high to make extraction cost effective. Extraction technologies and oil prices are global variables and are unlikely affected by local municipalities such as a county. Another threat to identification is a county that extracts a shale play to depletion, affecting other counties. Horizontally drilled wells typically have lateral length no longer than a couple of miles, and one county's well cannot reach another's shale formulation by a meaningful amount. The low permeability of shale formulations implies that shale wells in one county would not guzzle from a neighboring county's shale gas reservoir. Thus, shale plays influence the economic conditions of overlapping counties, but the counties would have minimal influence on each shale play's geological characteristics or resource availability.

1.4 Data on Environmental Belief

When identifying the effects of industry composition on environmental beliefs and policy, one challenge is that analysis requires individual-level data on environmentalism that can be matched to economic variables. A novel combination of data in this paper allows me to link individual policy opinions to an individual's employment status and geographic location, which can be mapped to local economic conditions.

1.4.1 Data on Environmental Belief and Vote Outcomes

Dependent variables of this paper, opinions on environmental policies and election outcomes, come from General Social Survey (GSS) data[10] provided by the NORC at the University of Chicago (NORC 2019). GSS is one of the longest-standing U.S.-wide surveys on individual socioeconomic status, policy opinions, and political behaviors. Since GSS contains information on individuals' environmental opinions, political beliefs, and reported vote outcomes in a single dataset, it provides an opportunity to study how economic conditions shape environmental beliefs and affect policy outcomes. GSS is based on personal interviews, minimizing measurement error in responses in comparison to other surveys that use mail, phone, or the Internet. On top of cross-section data from 1972, NORC constructed three biannual panel data in 2006, 2008, and 2010, each containing three biannual waves. To control for individual and year effects, I use panel data that covers 2006 to 2014.

I focus on variables that measure environmental beliefs, reported votes in past elections, and occupations of individuals and their families. To measure environmental beliefs, I use opinions on "improving and protecting the environment," "whether they think environmental scientists agree on global warming," and "whether they think environmental scientists understand global warming well." Individual reported votes are available for the 2004 and 2008 presidential elections. For those who did not or were unable to vote, hypothetical votes for 2004 and 2008 presidential candidates were surveyed. GSS data also provide the occupations of respondents and their family members. Using industry codes of respondents, spouses, fathers, and mothers, I observe whether a respondent's or a family member's employment in mining and related industries affected the respondent's environmental opinions.[11]

10. Some data used during analysis derived from Sensitive Data Files of the GSS, obtained under special contractual arrangements designed to protect respondents' anonymity. These data are not available from the author. People interested in obtaining GSS Sensitive Data Files should contact the GSS at GSS@NORC.org.

11. Survey questions are summarized in the appendix. More details on the questionnaire and the codebook

To match each respondent to the local economic conditions of the municipalities in which they reside, I acquired more refined geographic identifiers under contract with NORC. The public version of GSS contains all variables relevant to this study, but geographic locations of individuals are available only up to Census Division level. Since Census Divisions stretch over many states that have multiple shale plays in their boundaries, average economic conditions over these regions would be too broad and diluted. I thus narrow geographic units using sensitive data files that provide Census Tract-, county-, and state-level geographic identifiers. Since counties are sufficiently small that they are commonly contained in no more than one shale play, I match individual IDs to other variables at the county level. Most economic variables from other data are available up to the county level, making analysis feasible.

1.4.2 Methods to Link Economic Conditions to Environmental Belief

I use local economic conditions, or reliance on the mining industry, from a U.S. Census Bureau dataset. Based on shapefiles of shale and tight oil and gas provided by the EIA, I match shale geography to states and counties. I merge county-level economic variables, shale play geography, and previous data on environmental belief.

County-level economic activities in each industry are described in County Business Patterns (CBP), provided by the U.S. Census Bureau. CBP annually collects information on employment and payroll by North American Industry Classification System (NAICS) industry in each zip code, county, metropolitan area, congressional district, and state. Using employment in mining and drilling[12] relative to each county's total employment, I construct the primary explanatory variable, mining employment.

Shale play and basin geography come from shapefiles of shale gas and oil plays for the

---

are available on the NORC website (http://gss.norc.org/Get-Documentation/questionnaires)

12. I define the mining and drilling industry as all industries in the NAICS 21 category, which comprises extraction of natural resources, including minerals and fossil fuels. For brevity, I omit drilling and use the mining industry henceforth.

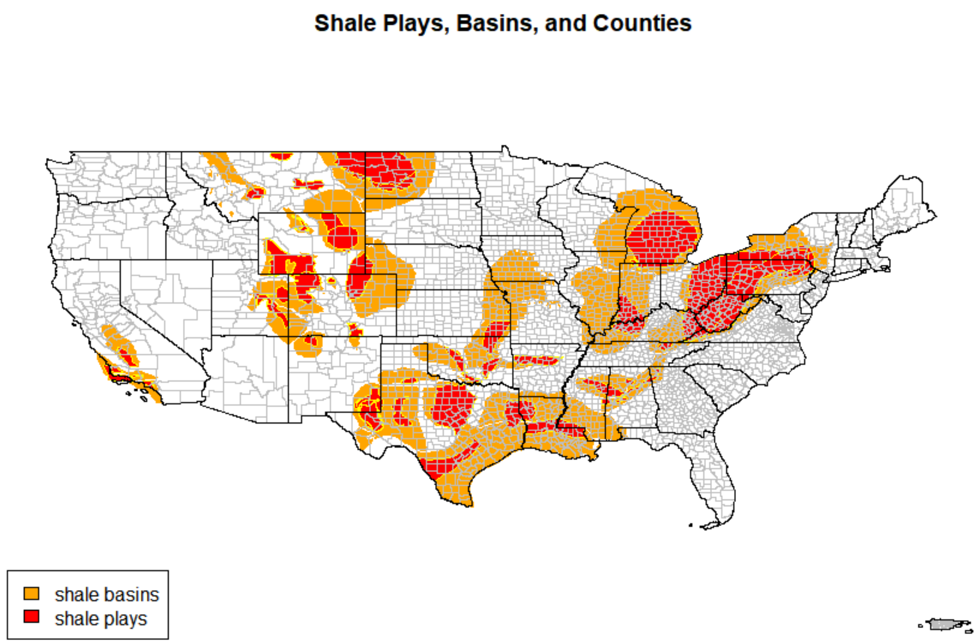


Figure 1.9: Shale Plays and Basins over Counties

Note: This figure plots shale gas and oil reservoirs in the lower 48 states. I overlay shale plays and basins on the boundaries of US counties. Shale basins are colored in orange and shale plays are colored in dark red. State boundaries are in thick black lines and county boundaries are in light grey lines.
Source: Shapefiles come from EIA (2016)

lower 48 states (EIA 2016)[13]. These maps contain information on each shale reservoir's boundaries and geologic characteristics. I overlay county-level boundaries onto shale boundaries using TIGER/Line Shapefiles from the U.S. Census Bureau (Census Bureau 2019a)[14]. For each shale play, I construct an indicator variable regarding whether a county is included in that shale play. I assign county-level mining employment and shale play dummy variables to each respondent in the GSS data. Resulting boundaries are shown in Figure 1.9.

1.4.3 Descriptive Statistics

Table 1.1 compares counties on shale plays (Column 2) and counties without shale resources (Column 1) in the combined GSS panel data (2006 – 2010). For this table and throughout the paper, counties with shale are defined as counties that have at least one shale play in its boundary. In my data, respondents in counties without shale and counties with shale represent 77.90% and 22.10%, respectively. For all types of counties, I report the mean and standard deviations (in parentheses) of belief measures, vote outcomes, economic conditions, and shale availability variables.

Panels A1 and A2 suggest that dependent variables, environmental beliefs and votes, were not significantly different in the data.

Panel B1 indicates that mining employment is about five times larger in counties on shale plays relative to the others. Although the difference is not statistically significant in the data, it is likely that counties on shale have relied systematically more on mining. However, Figure 1.8 suggests that mining employment was relatively parallel in both types of counties only until the fracking boom started in 2004.

Panel B2 shows that other economic variables are comparable as well. I proxy income with average wage, wealth with house price, and human capital by years of schooling. There are no significant differences in these variables.

13. Shapefiles can be downloaded from EIA's website (https://www.eia.gov/maps/maps.htm).

14. Shapefiles can be downloaded from the U.S. Census Bureau's website (https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html).

Table 1.1: Pre-trends and Levels, Counties with and without Shale

Summary of Beliefs, Votes, Economic Conditions, and Shale Resources, by County

	Without Shale	With Shale	All Counties
	Mean (SD)	Mean (SD)	Mean (SD)
Panel A. Dependent Variables			
Panel A1. Belief Measures			
Toomuch Envi	0.389 (0.488)	0.360 (0.480)	0.383 (0.486)
Sci Disagree	0.557 (0.497)	0.628 (0.484)	0.573 (0.495)
Sci Misunder	0.347 (0.476)	0.422 (0.495)	0.364 (0.481)
Panel A2. Vote Outcomes			
Vote for REP in 2008	0.395 (0.489)	0.382 (0.486)	0.392 (0.488)
Hypothetical Vote for REP	0.286 (0.452)	0.305 (0.461)	0.290 (0.454)
Panel B. Covariates			
Panel B1. Mining Industry			
% Employment in Mining	0.101 (0.447)	0.585 (1.666)	0.208 (0.900)
% Work in Mining (Narrow)	0.404 (6.346)	0.774 (8.768)	0.486 (6.955)
% Work in Mining	11.93 (32.41)	12.70 (33.30)	12.10 (32.61)
% Family Work in Mining (Narrow)	1.714 (12.98)	3.222 (17.66)	2.047 (14.16)
% Family Work in Mining	26.54 (44.16)	29.28 (45.51)	27.14 (44.47)
Panel B2. Other Economic Variables			
Avg Wage ($)	40712.4 (12062.7)	41025.2 (7437.4)	40781.5 (11206.6)
House price ($1K)	232.8 (146.7)	203.8 (132.6)	226.4 (144.2)
Years of Schooling	13.65 (3.073)	13.53 (2.969)	13.62 (3.051)
Panel C. Fracking Variables			
County has Shale	0 (0)	1 (0)	0.221 (0.415)
Num. Shale Plays in a County	0 (0)	1.497 (0.677)	0.331 (0.698)
New fossil value per capita ($)	73.02 (980.3)	321.5 (971.6)	127.9 (983.8)
Observations	11,379	3,228	14,607
Frequency in Sample	77.90 %	22.10 %	100 %

Note: This table compares counties with shale resources, counties without shale resource, and all counties in each column. I use shapefiles for US shale regions provided by EIA (2016) to determine if a county has shale resources. I compute the average and standard deviation of belief measures, vote results, mining employment, fracking variables, and other covariates in each panel. Standard errors are reported in parentheses.

In panel C, counties with shale plays have extracted four times more worth of fossil fuel. Whereas it indicates the impact of fracking on local economy, the differences are not significant. In the following section, I show that fracking indeed generated a substantial variation in mining employment, which lead to skepticism about environmental scientists and policies.

1.5 Estimation Framework

I construct a framework to estimate the influence of economic conditions, or mining employment, on environmental beliefs. I begin by using a simple linear model of belief and mining employment, and then explain threats to identification and fracking instrument variable for instrumental variable (IV) estimation.

1.5.1 Model of Environmental Belief and Mining Employment

Industry composition affects policy opinions and politics in several ways. This study quantifies the causal effect of additional mining employment on climate change opinions. The underlying mechanism is that due to job creation of mining firms, residents in areas that rely on mining industries have incentives to deny climate change. I provide measures of industry composition, environmental opinions, and policies, which I use to construct a linear estimation model:

$$Opinion_{i,t} = \beta_0 + \beta_1 Mining_{i,t} + \gamma_{s(i)} + \delta_t + u_{i,t} \tag{1.1}$$

$Opinion_{i,t}$ is a set of outcome variables that can be classified as belief measures and political opinion measures. Using NORC GSS data, I construct $Opinion_{i,t}$ for each year, t, according to each individual i's responses to questionnaires regarding their environmental beliefs and reported votes in previous presidential elections.

$Mining_{i,t}$ is a set of measures of the influence of the mining industry on local economies and each individual. The parameter of interest is β_1, which measures the causal influence of dependence on mining on environmental opinions.

I include state fixed effects $\gamma_{s(i)}$ to capture effects specific to each region. Since research suggests that the time trend of environmental opinions is significant, I include year fixed effects δ_t. $u_{i,t}$ is an error term, which might correlate with explanatory variable $Mining_{i,t}$.

In the remainder of this section, I explain how belief measures, election outcomes, and mining employment variables are constructed. I then explain how to treat endogeneity in $u_{i,t}$.

Constructing Environmental Belief Measures

Measures of environmental belief outcomes include

$$Opinion_{i,t} \in \{TooomuchEnvir_{i,t}, SciMisunder_{i,t}, andSciDisagree_{i,t}\}$$

These belief outcomes are binary variables constructed from categorical variables. The measure of skepticism regarding environmental protection, $ToomuchEnvir_{i,t}$, is based on raw variable $natenvir_{i,t}$, defined for each respondent, i,: [15]

$$natenvir_{i,t} = \begin{cases} 1 & \text{If i thinks we spend too little on environmental protection} \\ 2 & \text{If i thinks we spend about right on environmental protection} \\ 3 & \text{If i thinks we spend too much on environmental protection} \end{cases} \tag{1.2}$$

$SciMisunder_{i,t}$ is a measure of skepticism regarding scientists' understanding of climate change, constructed using raw variable $gwsci_{i,t}$, where:

15. The questionnaire and responses are described in more detail in the appendix. The survey records inapplicable responses 8. Don't know, 9. No answer, and 0. Not applicable. In the following description and during analysis, I remove inapplicable responses.

$$gwsci_{i,t} = \begin{cases} 1 & \text{If i thinks scientists understand global warming very well} \\ 5 & \text{If i thinks scientists understand global warming very poorly} \\ 2-4 & \text{In between, larger values mean i is more skeptical of scientists} \end{cases} \tag{1.3}$$

$SciDisagree_{i,t}$ describes whether respondents believe that most environmental scientists agree on global warming, constructed using raw variable $sciagrgw_{i,t}$ and defined as:

$$sciagrgw_{i,t} = \begin{cases} 1 & \text{If i thinks scientists totally agree with} \\ & \text{the existence and causes of global warming} \\ 5 & \text{If i thinks scientists disagree with global warming} \\ 2-4 & \text{In between, larger values mean i is more skeptical} \end{cases} \tag{1.4}$$

Although the raw variables carefully rank convictions or skepticism, it is unclear whether the categories represent the intensity of skepticism linearly. For interpretability and simplicity, I reduce these raw categorical variables into binary variables $ToomuchEnvir_{i,t}$, $SciMisunder_{i,t}$, and $SciDisagree_{i,t}$. I set the cutoffs at the median of the categorical variables and map all values less than the cutoff to be 0 and the others 1. Different cutoffs would result in different definitions of the resulting variables, but I avoid imposing an arbitrary structure.

Skepticism regarding environmental protection, $ToomuchEnvir_{i,t}$, is defined as:

$$ToomuchEnvir_{i,t} = \begin{cases} 0 & \text{if } natenvir_{i,t} = 1 \\ 1 & \text{if } natenvir_{i,t} = 2, 3 \end{cases} \tag{1.5}$$

$ToomuchEnvir_{i,t}$ is 0 if i thinks we spend too little on environmental protection and 1 if i thinks we should not spend more on environmental protection.

Skepticism regarding environmental scientists' ability to understand global warming, $SciMisunder_{i,t}$, is constructed as:

$$SciMisunder_{i,t} = \begin{cases} 0 & \text{if } gwsci_{i,t} = 1, 2 \\ 1 & \text{if } gwsci_{i,t} = 3, 4, 5 \end{cases} \tag{1.6}$$

$SciMisunder_{i,t}$ is 0 if i thinks scientists understand the presence and causes of global warming well enough, but 1 if i thinks scientists do not understand global warming well.

Skepticism regarding whether environmental scientists' view on global warming agrees is captured by $SciDisagree_{i,t}$:

$$SciDisagree_{i,t} = \begin{cases} 0 & \text{if } sciagrgw_{i,t} = 1, 2 \\ 1 & \text{if } sciagrgw_{i,t} = 3, 4, 5 \end{cases} \tag{1.7}$$

$SciDisagree_{i,t}$ is 0 if i is confident that scientists agree on global warming issues, but 1 if i is dubious about such agreement.

Election Outcome Variables

Election outcome variables are based on individual reported votes, but if respondents were unable to vote, hypothetical votes in previous presidential elections were captured instead. These variables are $\{Pres08_{i,t}, If08Who_{i,t}\}$. Using these variables, I construct Republican vote variable:

$$Opinion_{i,t} \in \{Rep08_{i,t+2}, IfRep08_{i,t}\}$$

.

Raw election variables $\{Pres08_{i,t}, If08Who_{i,t}\}$ are essentially dichotomous. Although survey questions included candidates outside of the Democratic and Republican parties, they comprise only two percent or less of responses. Taking $Pres08_{i,t}$ for example, possible responses were 1. Obama, 2. McCain, 3. Other candidate, 8. Don't know, 9. No answer, 0. Not applicable.[16] I drop responses other than 1. Democratic or 2. Republican.

$$Pres08_{i,t} = \begin{cases} 1 & \text{If i voted for Democrat in 2008} \\ 2 & \text{If i voted for Republican in 2008} \end{cases} \tag{1.8}$$

A significant portion of individuals did not or were unable to vote, but instead responded to the question on hypothetical votes. $If08Who_{i,t}$ documents hypothetical responses, for which answers are similar to those of $Pres08_{i,t}$.

$$If08Who_{i,t} = \begin{cases} 1 & \text{If i would have voted for Democrat in 2008} \\ 2 & \text{If i would have voted for Republican in 2008} \end{cases} \tag{1.9}$$

The first specification for the vote outcome is the two-year lead variable of the 2008 election:

$$Rep08_{i,t+2} = Pres08_{i,t+2} - 1 = \begin{cases} 0 & \text{If i will have voted for Democrat in 2008} \\ 1 & \text{If i will have voted for Republican in 2008} \end{cases} \tag{1.10}$$

I construct this variable to capture how current economic conditions affect individual political actions in the immediate future. I focus on 2008, during which two GSS panels (2006 and 2008) captured information regarding both 2008's economic conditions and respondents' votes during the 2008 election that were reported during 2010.

The second measure of political opinion is $IfRep08_{i,t}$, or willingness to vote for the

16. The questionnaire and responses are described in more detail in the appendix.

Republican candidate if the respondent could have voted during 2008.

$$If08Rep_{i,t} = If08Who_{i,t} - 1 = \begin{cases} 0 & \text{If i would have voted Democrat in 2008} \\ 1 & \text{If i would have voted Republican in 2008} \end{cases} \tag{1.11}$$

This variable captures respondents' retrospective evaluations of the 2008 presidential candidates. I use current, not forward, values of $If08Who_{i,t}$ since $If08Rep_{i,t}$ reflects the current sentiment on candidates, where $Rep08_{i,t+2}$ must be a lead variable since it represents past actions that cannot be affected by the current economic situation.

Mining Employment and Occupation Variables

I measure the influence of the mining industry, $Mining_{i,t}$, at the county aggregate- and individual levels. The first measure is the portion of the mining industry in each county's employment, $MiningEmployment_{c(i),t}$, which captures how the dependence of county c on the mining industry affects individual i's beliefs, where $c = c(i)$ is the county in which i resides. The other variables, $MiningSelf_{i,t}$ and $MiningFamily_{i,t}$, describe whether respondent i is employed in the mining sector and whether i has a family member employed in the mining sector, respectively.

$MiningEmployment_{c(i),t}$ is constructed using employment data from County Business Patterns (CBP), which contains the number of employees in each NAICS industry, $NumberofEmployeesinIndustryX_{c,t}$. I set $MiningEmployment_{c(i),t}$ to be the portion of employment in the mining and drilling industry in each county.

$$MiningEmployment_{c,t} = \frac{Number\ of\ Employees\ in\ Mining_{c,t}}{Number\ of\ Employees\ in\ All\ Sectors_{c,t}} \tag{1.12}$$

GSS data also provide the occupations of respondents and their family members. Based on the industry code of self, spouse, father, and mother, I observe whether being employed or having a family member employed in mining and related industries affects respondents' environmental opinions. Raw variables in the GSS data are:

$$\begin{cases} SelfIndus_{i,t} = x \text{ if i is employed in industry code } x \\ SpouseIndus_{i,t} = x \text{ if i's spouse is employed in industry code } x \\ FatherIndus_{i,t} = x \text{ if i's father is employed in industry code } x \\ MotherIndus_{i,t} = x \text{ if i's mother is employed in industry code } x \end{cases} \tag{1.13}$$

Using these variables, I define a set of industries that are mining or related:

$$MiningRelated = \{x|370 \leq x \leq 490 \text{ or } 2070 \leq x \leq 2090\}$$

and construct indicator variables regarding whether a respondent or family member is employed in mining or related sectors:

$$MiningSelf_{i,t} = \begin{cases} 1 & \text{if } SelfIndus_{i,t} \in MiningRelated \\ 0 & \text{otherwise} \end{cases} \tag{1.14}$$

$$MiningSpouse_{i,t} = \begin{cases} 1 & \text{if } SpouseIndus_{i,t} \in MiningRelated \\ 0 & \text{otherwise} \end{cases} \tag{1.15}$$

$$MiningFather_{i,t} = \begin{cases} 1 & \text{if } FatherIndus_{i,t} \in MiningRelated \\ 0 & \text{otherwise} \end{cases} \tag{1.16}$$

$$MiningMother_{i,t} = \begin{cases} 1 & \text{if } MotherIndus_{i,t} \in MiningRelated \\ 0 & \text{otherwise} \end{cases} \tag{1.17}$$

I define an indicator that i has a family member employed in mining or related sectors:

$$MiningFamily_{i,t} = max\{MiningSpouse_{i,t}, MiningFather_{i,t}, MiningMother_{i,t}\} \quad (1.18)$$

During analysis, I use $MiningSelf_{i,t}$ and $MiningFamily_{i,t}$ to measure the influence being employed or having a family member employed in the mining sector has on environmental beliefs.

1.5.2 Sources of Endogeneity and Fracking Instruments

Two sources of endogeneity are present in the main model (1.1)—reverse causality and unobserved factors. I treat reverse causality using fracking instruments and use state fixed effects to control for time-invariant unobservables.

Sources of Endogeneity

One of the main challenges in identification is that not only the environmental belief is a function of mining employment, but also the reverse could be true. Measures of mining industry's impact on local economy, $Mining_i, t$, are determined by a collection of mining firms' decisions to enter a region, extract resources, and hire employees. The firms' decisions, in turn, are mainly based on two factors: First, only the regions that natural resources are geologically present will be considered by the firms. Then, the firm evaluates the profitability of extracting in each region. This profitability is a function of variables including global resource price, infrastructure, labor force, and local regulations. Although the geological availability of resources is pre-determined and plausibly unaffected by environmental beliefs, the environmental opinion could have a direct impact on regulation on mining and, in turn, repel mining firms. If mining firms disproportionately enter areas favorable to them, OLS will overestimate the impact of the mining dependence $Mining_i, t$ on the environmental belief.

Fracking Instrument: Shale play-by-year Fixed Effects

As technological shocks accumulate, some shale resources that were previously inaccessible become extractable. In order to capture the variation in the profitability of shale play due to the geographic characteristics and technological advances that allows for extraction, I construct shale play (p)-by-year (t) indicator variables for each county c.

$$\mu_{p,t} = \mu_{p(c),t} = \sum_{\tilde{p} \in P} \mu_{\tilde{p},t} \times 1\{c \in \tilde{p}\} \tag{1.19}$$

The presence of shale in a county and the geography of each shale play come from the US Energy Information Administration. I use the shale play-by-year fixed effect to set up a two-stage regression equation.

$$Mining_{i,t} = \sum_{p \in P} \mu_{p,t} \times 1\{c(i) \in p\} + g_c + d_t + e_{c,t} \tag{1.20}$$

From the first stage, I compute the predicted value of $Mining_{i,t}$

$$\hat{Mining}_{i,t} = \sum_{p \in P} \hat{\mu_{p,t}} \times 1\{c(i) \in p\} + \hat{g_c} + \hat{d_t} \tag{1.21}$$

and finally estimate the following equation

$$Opinion_{i,t} = \beta_0 + \beta_1 \hat{Mining}_{i,t} + \gamma_{s(i)} + \delta_t + u_{i,t} \tag{1.22}$$

I exploit the exogenous variation generated by new gas and oil production from the shale play the sample counties are part of; this variation is outside the control of each county and gives a credible treatment shock of mining employment to counties that are otherwise similar. Consider a toy example where county a is on shale play p_a and county b is on a different shale play p_b. Assume that both counties have similar characteristics but only the depth of shale plays differ; p_a is 5,000 feet deep and p_b is 10,000 feet deep. If fracking technology and global oil price makes it profitable to extract up to 6,000 feet in 2006 but up to 12,000 feet in 2008, then county b becomes eligible for fracking but county a faces the same eligibility. Assuming there are no defiers in the sample and the above change in eligibility is not driven by environmental beliefs, this shale play-by-year fixed effect is an exogenous instrument variable for mining employment.

This setup also alleviates issues with unobservables by controlling for state fixed effects $\gamma_{s(i)}$ and year fixed effects δ_t. Whereas local history or culture could shape environmental belief and there might be time trends in environmental opinions,[17] any time invariant variables specific to a region and any year effects are controlled under this specification.

1.6 Findings

Using exogenous shocks to the local economy, I find that mining employment affects environmental beliefs and policy. The portion of mining in total local employment (1) reduces support for environmental protection and reduces trust in climate scientists and (2) increases support for republican candidates during elections, and (3) being employed or having a family employed in the mining industry reduces support for environmental protection.

1.6.1 Local Mining Employment and Environmental Belief

I estimate the main equation (1.1) following the two-step procedure (1.20)-(1.22). Table 1.2 summarizes the impact of mining employment on the first measure of environmental belief, opinions on the environmental protection.

The binary dependent variable $ToomuchEnvir_{i,t}$ is 1 if the respondent thinks we spend too much on the environmental protection. Columns (1)-(4) instruments on the shale play-by-year fixed effect, and Columns (5)-(8) describe the OLS results. In Columns (1) and (5), I do not control for regional fixed effects, where I put state, commuting zone, or county

17. Kahn and Kotchen (2011), for example, demonstrates that people are more skeptical in years under recession.

Table 1.2: Mining Employment and Opinions on Environmental Protection

Dep. Var: We spend too much on environment (ToomuchEnvir)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Employment in Mining	0.026**	0.039***	0.239	0.000	0.021***	0.019**	0.013	0.000
	(0.012)	(0.012)	(0.213)	(.)	(0.007)	(0.008)	(0.018)	(0.023)
Constant	0.378***				0.379***	0.379***	0.381***	0.384***
	(0.006)				(0.006)	(0.006)	(0.007)	(0.007)
Method	IV	IV	IV	IV	OLS	OLS	OLS	OLS
Region FE	none	state	czone	county	none	state	czone	county
Obs	7055	7050	6932	6960	7055	7050	6932	6960
Dep var mean	0.383	0.383	0.383	0.384	0.383	0.383	0.383	0.384

Note: Dependent variable, "ToomuchEnvir", is a binary variable that measures respondents' skepticism on environmental protection. This variable equals 1 if the respondent thinks we should not spend more on protecting the environment and 0 if they think we should spend more. Columns 1-4 are based on IV regression and columns 5-8 report the OLS results. Columns 2 and 6 include state fixed effects, columns 3 and 7 include commuting zone fixed effects, and columns 4 and 8 include county fixed effects. Since GSS is a short panel, commuting zone fixed effects and county fixed effects significantly reduce degrees of freedom. The preferred specification is column 2, which is an IV estimation result that controls for time-invariant regional effects on state level. * p<.10, ** p<.05, *** p<.01. Standard errors in parentheses.

controls for the rest of the Columns.

The preferred specification is Column (2), IV regression with state fixed effect, as it estimates causal effect without severly suffering from multicolinearity or omitted variable bias. Commuting zones and counties are more refined geographic units and would better capture municipality-specific effects, but these geographies and shale play boundaries are heavily multicolinear. For instance, the standard error for the explanatory variable increases from 0.012 to 0.213 in Columns (2) and (3). Furthermore, shale plays and counties are perfectly colinear, making IV regression with county fixed effects infeasible. On the other hand, excluding any regional fixed effects might lead to omitted variable bias as different states are likely differ in legislation, regulation, and culture. Hence, state fixed effect captures a necessary degree of confounding effects without sacrificing the feasibility of the regression.

Table 1.3 reports the estimated effect of mining employment on one's skepticism about the agreement of environmental scientists on global warming. The specifications are analogous to those of Table 1.2, and I focus on Column (2). Tables 1.2 and 1.3 show how environmental opinions change as mining industry employs additional one percentage point of each county's total employment. On average, it leads to a resident's support for environmental protection

Table 1.3: Mining Employment and Skepticism about Agreement between Scientists

Dep. Var: Scientists disagree on global warming (SciDisagree)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Employment in Mining	0.007	0.042	0.569	0.000	-0.012	0.024	-0.097	-0.182*
	(0.028)	(0.036)	(1.064)	(.)	(0.018)	(0.030)	(0.080)	(0.110)
Constant	0.571***				0.575***	0.568***	0.592***	0.611***
	(0.014)				(0.013)	(0.014)	(0.021)	(0.026)
Method	IV	IV	IV	IV	OLS	OLS	OLS	OLS
Region FE	none	state	czone	county	none	state	czone	county
Obs	1463	1461	1434	1427	1463	1461	1434	1427
Dep var mean	0.573	0.573	0.573	0.575	0.573	0.573	0.573	0.575

Note: Dependent variable, "SciDisagree", is a binary variable that measures respondents' skepticism on the agreement between climate scientists. This variable equals 1 if the respondent thinks scientists remain split on climate change issues and 0 if they think climate scientists agree. Columns 1-4 are based on IV regression and columns 5-8 report the OLS results. Columns 2 and 6 include state fixed effects, columns 3 and 7 include commuting zone fixed effects, and columns 4 and 8 include county fixed effects. Since GSS is a short panel, commuting zone fixed effects and county fixed effects significantly reduce degrees of freedom. The preferred specification is column 2, which is an IV estimation result that controls for time-invariant regional effects on state level. * p<.10, ** p<.05, *** p<.01. Standard errors in parentheses.

to decrease by 3.9 percentage points and one's skepticism about climate scientists to increase by 4.2 percentage points. Although these coefficients exhibit the same sign, the effect on $ToomuchEnvir$ is statistically significant where the coefficient for $SciDisagree$ is not significantly different from zero. A potential reason is that $SciDisagree$ variable is observed only in 2006 and 2010, and are surveyed only in a subset of the respondents even in those years. As a result, the sample size is smaller relative to that of $ToomuchEnvir$. Hence, I note that the estimates on the skepticism on the scientists should be interpreted with caution.

1.6.2 Occupation in Mining and Environmental Belief

Table 1.4 summarizes how being employed or having a family member employed in the mining industry reduces support for environmental protection. Similarly, Table 1.5 provides results on the impact of occupation of self and family on one's skepticism on environmental scientists.

Columns (1) and (3) report relationship between being employed in mining sector and one's opinion on environmental protection. The explanatory variable in these Columns, $MiningSelf$, is binary and equals 1 if the respondent is employed in mining or related

Table 1.4: Occupation and Opinions on Environmental Protection

Dep. Var: We spend too much on environment (ToomuchEnvir)

	(1)	(2)	(3)	(4)
Work in Mining (Self)	0.527**		0.032*	
	(0.252)		(0.018)	
Work in Mining (Family)		0.202		0.015
		(0.174)		(0.013)
Constant			0.379***	0.379***
			(0.006)	(0.007)
Method	IV	IV	OLS	OLS
Region FE	state	state	state	state
Obs	7050	7050	7050	7050
Dep var mean	0.383	0.383	0.383	0.383

Note: Dependent variable, "ToomuchEnvir", is a binary variable that measures respondents' skepticism on environmental protection. This variable equals 1 if the respondent thinks we should not spend more on protecting the environment and 0 if they think we should spend more. Columns 1 and 2 are based on IV regression and columns 3 and 4 report the OLS results for reference. Column 1 reports the impact of working in mining sector on one's environmental belief. Column 2 reports the impact of having a family working in mining sector on one's environmental belief. I include state fixed effects in all specifications. * p<.10, ** p<.05, *** p<.01. Standard errors in parentheses.

sector. Columns (2) and (4) documents the relationship between having a family member working in mining sector, or $MiningFamily = 1$, and one's perception about environmental protection. [18]

Based on the discussion above, I present results only with state fixed effects in these table. For comparison I provide OLS results in Columns (3) and (4), but the main specifications are Columns (1) and (2).

Column (1) of Table 1.4 shows that working in mining sector makes a respondent 52.7 percentage point more reluctant to spend more on environmental protection. This is significant at 5 percent level. Column (2) of Table 1.4 indicate that family's occupation could affect support on environmental protection by 20.2 percentage point, but this is statistically insignificant and should not be seen as a strong evidence.

Columns (1) and (2) of Table 1.5 show that one's occupation also has ambiguous impact on one's skepticism about the environmental scientists. Similar to the case in Table 1.3, small sample size would have lead to this inconclusive results.

18. See equations (1.14) and (1.18) for the definitions.

Table 1.5: Occupation and Skepticism about Agreement between Scientists

Dep. Var: Scientists disagree on global warming (SciDisagree)

	(1)	(2)	(3)	(4)
Work in Mining (Self)	-0.619		-0.015	
	(0.617)		(0.041)	
Work in Mining (Family)		-0.504		0.057*
		(0.424)		(0.030)
Constant			0.575***	0.558***
			(0.014)	(0.015)
Method	IV	IV	OLS	OLS
Region FE	state	state	state	state
Obs	1461	1461	1461	1461
Dep var mean	0.573	0.573	0.573	0.573

Note: Dependent variable, "SciDisagree", is a binary variable that measures respondents' skepticism on the agreement between climate scientists. This variable equals 1 if the respondent thinks scientists remain split on climate change issues and 0 if they think climate scientists agree. Columns 1 and 2 are based on IV regression and columns 3 and 4 report the OLS results for reference. Column 1 reports the impact of working in mining sector on one's environmental belief. Column 2 reports the impact of having a family working in mining sector on one's environmental belief. I include state fixed effects in all specifications. * p<.10, ** p<.05, *** p<.01. Standard errors in parentheses.

1.6.3 Local Mining Employment and Vote Outcomes

Mining employment also increased votes for Republican candidates in the 2008 presidential election. Column (1) of Table 1.6 shows that if mining industry employs additional one percentage point from the total county employment in year 2008, the vote for Republican in that county increased by 3.2 percentage points at the presidential election later in that year.

Although the point estimates of the impact of occupation are substantial (Columns 2-3), they are not significant at 5 percent level. Compared to OLS results in Columns (5) and (6), the lack of power would have stemmed from the potential colinearity of play-by-year fixed effects and county fixed effects.

Results on hypothetical votes are summarized in Table 1.7. The results are noisy due to two potential reasons: small sample and the fact that respondents might have answered more casually in $If08Rep$ than in $Pres08$ as it surveys one's intention rather than one's actual political behavior.

Table 1.6: Mining Industry and Vote Outcomes

Dep. Var: Votes for Republican (Pres08)

	(1)	(2)	(3)	(4)	(5)	(6)
% Employment in Mining	0.032***			0.030***		
	(0.011)			(0.010)		
Work in Mining (Self)		0.560			0.068***	
		(0.424)			(0.023)	
Work in Mining (Family)			0.380*			0.055***
			(0.200)			(0.016)
Constant				0.389***	0.387***	0.380***
				(0.007)	(0.007)	(0.008)
Method	IV	IV	IV	OLS	OLS	OLS
Region FE	state	state	state	state	state	state
Obs	4813	4813	4813	4813	4813	4813
Dep var mean	0.395	0.395	0.395	0.395	0.395	0.395

Note: Dependent variable, "Pres08", indicates whether the respondent voted for a Republican (=1) or a Democrat (=0) candidate in 2008 presidential election. Columns 1 - 3 are based on IV regression and columns 4 - 6 report the OLS results for reference. Column 1 reports the changes in one's votes when mining industry employs additional one percentage of worker in the county they reside. Column 2 reports the impact of working in mining sector on one's vote. Column 3 reports the impact of having a family working in mining sector on one's vote. I include state fixed effects in all specifications. * p<.10, ** p<.05, *** p<.01. Standard errors in parentheses.

Table 1.7: Mining Industry and Hypothetical Vote Outcomes

Dep. Var: Hypothetical Votes for Republican (If08Rep)

	(1)	(2)	(3)	(4)	(5)	(6)
% Employment in Mining	0.009			0.015		
	(0.012)			(0.011)		
Work in Mining (Self)		-0.875			0.057*	
		(0.577)			(0.031)	
Work in Mining (Family)			0.045			0.053**
			(0.256)			(0.024)
Constant				0.286***	0.282***	0.277***
				(0.011)	(0.011)	(0.012)
Method	IV	IV	IV	OLS	OLS	OLS
Region FE	state	state	state	state	state	state
Obs	1789	1789	1789	1789	1789	1789
Dep var mean	0.290	0.290	0.290	0.290	0.290	0.290

Note: Dependent variable, "If08Rep", indicates whether the respondent would have voted for a Republican (=1) or a Democrat (=0) candidate in 2008 presidential election. Columns 1 - 3 are based on IV regression and columns 4 - 6 report the OLS results for reference. Column 1 reports the changes in one's hypothetical votes when mining industry employs additional one percentage of worker in the county they reside. Column 2 reports the impact of working in mining sector on one's hypothetical vote. Column 3 reports the impact of having a family working in mining sector on one's hypothetical vote. I include state fixed effects in all specifications. * p<.10, ** p<.05, *** p<.01. Standard errors in parentheses.

1.7 Conclusions

I find that mining employment may influence people to be skeptical of environmental scientists and reduce their support for environmental policies. Results of this study contribute to extant literature by using plausibly exogenous shocks to assess a causal link.

Findings show that the fracking boom substantially influenced environmental opinions and elections. This proposes to consider that local industry composition and economic conditions when studying the political economy. Despite the increasing evidences that resource extraction activities directly affect welfare of the residents, the literature remains silent on whether the resource boom also intervene in policy decision process.

However, the exact channel through which mining firms change policy opinions remains unknown. Future research should explore one possible explanation—media choice could be affected by the existence of mining and fracking. Also, adopting alternative instruments would further alleviate concerns that the fracking boom could also influence the environmental belief in channels other than mining employment. I suggest in future research to exploit the data on coal mine code violations, provided by Mine Safety and Health Administration (MSHA). A careful comparison of mine employment before and after code violation happens in a municipality might generate an extra source of exogenous variation in mining employment.

CHAPTER 2

NOTCHES IN ELECTRICITY MARKETS

2.1 Introduction

A principle in standard economics states that rational agents optimize at the margin, but recent empirical studies suggest that consumers do not internalize nonlinear price structure fully. Under standard assumptions, the theory predicts the bunching behavior—an excess distribution of consumers—around points of discontinuities when marginal prices or total payments jump. Nonetheless, consumers experiencing nonlinear prices might respond to alternative price measures rather than marginal prices. Research suggests that consumers respond to expected marginal prices (Borenstein 2009) or average prices (Ito 2014; Liebman and Zeckhauser 2004), or overreact to kinks (Shaffer 2019).

In many labor and consumption goods markets, two types of nonlinear prices are common—kink points and notch points. Kinks refer to discontinuities in slopes of total prices, or jumps in marginal prices, and notches refer to discrete jumps in total price. Common examples of kink points are found in tax schemes in most countries, including the U.S. Federal Income Tax that imposes disparate tax rates within income brackets. The marginal tax rate increases past each kink point, and taxpayers experience continuous, monotonic, and piecewise linear budget lines. A seller or regulator might impose lump-sum discounts or surcharges for purchases that exceed some amount, or a notch point. Agents thus experience discontinuities in their budget sets in addition to potential slope variation. In terms of bunching incentives, notches are more salient and responses to discontinuities more dramatic.

Notches are pivotal tools to studying agents' behaviors, and are prevalent in various applications. However, the literature has focused on taxation and labor markets, and thus less is known about how consumers respond to notch points in consumption goods markets. Bunching analyses of kinks and notches allow structural estimation of elasticities, but notches also contain information about optimization frictions. Although both kinks and notches are studied extensively in labor market contexts, research on nonlinear prices in consumption goods markets seldom involves jumps in total payments or budget lines.

To study whether consumers respond to notches in consumption goods market, I exploit a nonlinear tariff and a major reform in Korea's residential electricity market. Unlike in most other residential electricity markets, jumps in total payments are sizable and a substantial portion of consumers lie around these notch points. Marginal prices, the slopes of total payments, change significantly around these notch points. This context thus provides an opportunity to study a consumption goods market with substantial jumps in both the slope and level of budget lines.

I test whether consumers respond to prices that are easier to understand, and whether they fail to internalize the full price structure. I estimate demand elasticities to marginal, naive marginal, and average prices. A consumer responds to naive marginal prices if he/she understands that marginal prices are increasing but do not account for jumps in total prices and both discounts and surcharges in special situations. Hence, naive marginal prices represent more precise measures than average prices do, but they are rougher measures than marginal prices.

I show that consumers mostly respond to average prices, and marginal and naive marginal prices do not explain demand properly. Estimated average price elasticity is consistent with the literature. However, marginal and naive marginal prices' elasticities are positive when controlling for average prices, a behavior that accords with consumers' bunching behavior. Economic theory predicts that customers bunch before points of nonlinearity, but I observe bunching after these notch and kink points. Lagged responses do not drive this anomaly, since previous months' prices have no influence on demand.

This paper contributes to nonlinear pricing literature by providing insights into notches outside of labor markets. Despite their importance, notches are seldom studied in consumption goods contexts. For example, Kleven and Waseem (2013) analyzes notches in labor markets. Notches in consumption goods and labor markets impose qualitatively different nonlinear structures, and consumers are less aware of price structures in the former. Hence, bunching analyses in labor markets cannot be extrapolated to other markets without adjustment.

Although notches are rarer, nonlinearities caused by kinks are studied commonly in electricity and other markets. Borenstein (2009) and Ito (2014) proposed that consumers in residential electricity markets do not optimize at the margin under nonlinearity. Not only do total payments jump, the tariff is highly nonlinear; the lowest and highest marginal prices differ by up to 11.7 times.

I add to consumer misperception literature that finds that consumers who experience nonlinear prices fail to optimize due to behavioral reasons. Sallee (2014) suggests that rational inattention hinders consumers from perceiving correct marginal prices. Using an experiment, Kahn and Wolak (2013) finds that providing an explanation of a price scheme resolves inattention. McRae and Meeks (2016) argues that some households understand nonlinear electricity tariffs better than others do. Evidence that information provision influences consumer behavior is found in house weatherization Allcott and Greenstone (2017), automobile Allcott and Knittel (2019), light bulb Allcott and Taubinsky (2015), electronic appliances Hausman (1979), and financial asset Carvalho and Silverman (2019) markets. Nonetheless, studies, including Busse, Knittel, and Zettelmeyer (2013), suggest that consumers' behaviors lie within a theoretically reasonable range.

2.2 Theory of Notch Points and Bunching

Notches refer to jumps in total prices, often caused by lump-sum surcharges or discounts on purchases after certain points. In labor markets, theory and data suggest that notches generate bunching behaviors, followed by a collapse in density. However, there are differences in how notches influence consumers' budget lines and choices in consumption goods contexts. I provide a theory of notches in consumption goods markets that explains the consequences of discrete jumps in labor and electricity markets.

2.2.1 *Consumer Preferences*

Consumers have a quasi-linear, iso-elastic utility functionL

$$u = c + \frac{n}{1+1/e}(\frac{x}{n})^{1+1/e}, \tag{2.1}$$

where c represents consumption (other than electricity) and x represents electricity consumption. n is a preference parameter for electricity such that the optimal x increases in n, and e is the price elasticity of electricity demand.

The consumer experiences budget set:

$$c + P(x) = L \tag{2.2}$$

Consumers experience fixed income L and total payment scheme $P(x)$ for their electricity usage x. In the case of linear prices, $P(x) = p \cdot x$, where p is the constant marginal price of electricity. With notch point $x*$, total payment becomes:

$$P(x) = p \cdot x + 1\{x > x*\}[\Delta P + \Delta p \cdot x] \tag{2.3}$$

Two sources generate notch point $x*$—a lump sum surcharge $\Delta P > 0$ (or discount if $\Delta P < 0$) and a jump in payment slope Δp. My data suggests a positive jump in prices, and I focus on the case $\Delta p > 0$ in this section.

I assume a smooth distribution of parameter n, and denote the cumulative distribution function by $F(n)$ and the density by $f(n)$. For the linear case, the utility function becomes:

$$u(x) = [L - p \cdot x] + \frac{n}{1+1/e}(\frac{x}{n})^{1+1/e} \tag{2.4}$$

and a first-order condition solves the constraint optimization problem:

$$u'(x) = -p + (\frac{x}{n})^{1/e} = 0 \tag{2.5}$$

$$x = n \cdot p^e \tag{2.6}$$

Interior solutions for nonlinear cases are similar, where p is substituted with marginal prices p or $p + \delta p$, depending on the location of the solution.

This first-order condition suggests important characteristics of parameters e and n. Equation (2.6) corroborates the definition of elasticity e:

$$\frac{dx/dp}{x/p} = (n \cdot e \cdot p^{e-1}) \cdot p/(n \cdot p^e) = e \tag{2.7}$$

The optimal x in equation (2.6) also shows that electricity consumption is not only increasing in but also proportional to n. Another observation is the limiting case of Leontieff prefernces when $e \to 0$:

$$\lim_{e \to 0} x = n \tag{2.8}$$

Similarly, when $p = 1$, electricity consumption equals parameter n. Combining these observations, n represents a benchmark of electricity consumption.

Marginal Consumer at the Notch Point

Using the results above, I describe the behavior of a marginal consumer who is indifferent between notch point $x*$ and an interior point. Combining the consumer's responses and the distribution around the notches, I express elasticity as an implicit function of bunching responses and price jumps.

I assume a single notch point, and I discuss the case of multiple discontinuity points later. For simplicity, I suppose that notch $x*$ is introduced to an originally linear scheme with price $P(x) = p \cdot x$. Recall the heterogeneity of parameter n $F(n)$ and relationship $x = n \cdot p^e$, which show that electricity consumption x follows a smooth distribution $H(x)$ in interior points, or without notch points.

$$H(x) = F(x/p^e)$$
$$h(x) = f(x/p^e)/p^e$$

Now assume that additional marginal and total payments, Δp and ΔP, are introduced for points $x > x*$. Under reasonable assumptions, smooth distribution $F(n)$ implies a well-defined range $[n_L, n_H]$ that consumers choose to consume at the notch point. For consumers with $n = n_L$, notch point $x*$ was also the interior solution when prices were linear. The highest type consumers with n_H are indifferent between notch point $x*$ and interior point x^I, hence the marginal consumers. Denote their preference parameter as $n_H = n * + \Delta n *$, and let $x * + \Delta x *$ be their original consumption before the notch was introduced.

Let the utility of consumer n_H at the interior point be $u^I = u(x^I)$ and at the notch be $u^N = u(x*)$. By n_H's definition, $u^N = u^I$ must hold. It is also useful to derive the following relationship from the first-order condition:

$$x^I = (n * + \Delta n *)(p + \Delta p)^e \tag{2.9}$$

From the utility and electricity consumption before the notch:

$$(n * + \Delta n *) = (x * + \Delta x *)/p^e \tag{2.10}$$

I combine these observations to solve elasticity for the function of the price structure and consumer responses. I equate the two utilities:

$$\begin{aligned} u^N &= L - p \cdot x * + \frac{n * + \Delta n *}{1 + 1/e}\left(\frac{x*}{n * + \Delta n *}\right)^{1+1/e} \\ = u^I &= L - \Delta P - (p + \Delta p)x^I + \frac{n * + \Delta n *}{1 + 1/e}\left(\frac{x^I}{n * + \Delta n *}\right)^{1+1/e} \end{aligned} \tag{2.11}$$

Using equation (2.9) to substitute x^I in:

$$\begin{aligned} u^I &= L - \Delta P - (p+\Delta p)(n*+\Delta n*)(p+\Delta p)^e + \frac{n*+\Delta n*}{1+1/e}(\frac{(n*+\Delta n*)(p+\Delta p)^e}{n*+\Delta n*})^{1+1/e} \\ &= L - \Delta P - (n*+\Delta n*)(p+\Delta p)^{1+e} + \frac{n*+\Delta n*}{1+1/e}(p+\Delta p)^{1+e} \end{aligned} \tag{2.12}$$

Continuing to solve for equation (2.12):

$$\begin{aligned} 0 =& u^N - u^I \\ =& L - p\cdot x* + \frac{n*+\Delta n*}{1+1/e}(\frac{x*}{n*+\Delta n*})^{1+1/e} \\ & - [L - \Delta P - (n*+\Delta n*)(p+\Delta p)^{1+e} + \frac{n*+\Delta n*}{1+1/e}(p+\Delta p)^{1+e}] \\ =& \Delta P - p\cdot x* + \frac{n*+\Delta n*}{1+1/e}(\frac{x*}{n*+\Delta n*})^{1+1/e} + (n*+\Delta n*)\frac{1}{1+e}(p+\Delta p)^{1+e} \end{aligned} \tag{2.13}$$

Substitute $n*+\Delta n*$ terms using equation (2.10):

$$\begin{aligned} 0 &= u^N - u^I \\ &= \Delta P - p\cdot x* + \frac{x*+\Delta x*}{1+1/e}\frac{1}{p^e}(\frac{x*}{x*+\Delta x*}p^e)^{1+1/e} + (x*+\Delta x*)\frac{1}{p^e}\frac{1}{1+e}(p+\Delta p)^{1+e} \end{aligned} \tag{2.14}$$

Divide both sides by $(x*+\Delta x*)p$ to obtain the final equation:

$$\frac{1}{1+\Delta x*/x*}[1-\frac{\Delta P/x*}{p}] - \frac{1}{1+1/e}[\frac{1}{1+\Delta x*/x*}]^{1+1/e} - \frac{1}{1+e}[1+\frac{\Delta p}{p}]^{1+e} = 0 \tag{2.15}$$

Equation (2.15) expresses elasticity e as an implicit function of the percentage of electricity consumption responses $\Delta x*/x*$ and changes to average price, $\frac{\Delta P/x*}{p}$ and $\frac{\Delta p}{p}$ that are, respectively, generated by ΔP and Δp. Although there is no closed form solution, it is possible to solve for e given estimates of the other arguments.

Before discussing estimation, I point out that price schemes' multiple discontinuity points can be treated locally as a single notch case. In this case, the steps similar to Kleven and Waseem (2013) can be followed.

Estimation Procedure

Since price elasticity e is expressed as a function, I introduce steps to estimate the elasticity and other arguments of the implicit function. Essential is determining the original density without the notch point, computing the size of bunching, and using the relationship between the electricity consumption responses and the size of bunching.

Excess bunching, B, can be approximated using electricity consumption responses $\Delta x*$:

$$B = \int_{x*}^{x*+\Delta x*} h(x)dx \approx h(x*)\Delta x*, \tag{2.16}$$

where $h(x)$ is the original density of x prior to the notch's introduction. This approximation assumes that the density is close to constant in the interval $[x*, x*+\Delta x*]$.

Next is estimating density $h(x)$ as if there were no notches. The insight is that points far from the notch point reveal the shape of the density, but the neighborhood of $x*$ is affected by nonlinearity and should be excluded. Later, I discuss how to obtain optimal region $[x_L, x_U]$, but assume that this region is given for now. Partition the support of x into smaller bins with index j, exclude the range $[x_L, x_U]$, and then fit the histogram of observed x with a polynomial of degree p. For each j, denote the polynomial as:

$$\hat{h}_j = \sum_{k=0}^{p} \hat{\beta}_j \cdot (z_j)^k \tag{2.17}$$

Using this fitted density polynomial, compute excess bunching as the sum of excess counts in the region between x_L and $x*$:

$$\hat{B} = \sum_{\{j:x_L \leq x_j \leq x*\}} (h_j - \hat{h}_j) \tag{2.18}$$

Similarly, missing mass is:

$$\hat{M} = \sum_{\{j: x* \leq x_j \leq x_U\}} (\hat{h}_j - h_j) \tag{2.19}$$

Since this process depends on region $[x_L, x_U]$, it is important to determine lower and upper bounds precisely. The theory suggests that excess bunching is so sharp that it is possible to determine x_L in the histogram visually.

The region above the notch, x_U, can be determined automatically under a plausible assumption. Suppose that the notches do not permute the order of n among customers. This effectively implies that the marginal buncher is well-defined and rules out the possibility that consumers previously far beyond x_U start to bunch at $x*$. Under this assumption, it must hold that $\hat{M} = \hat{B}$. Also, $\hat{M}$ is increasing in x_U, as long as $\hat{h}_j > h_j$ in that region. Hence, $\hat{B}$ can be computed. Start with a very low $x_U \approx x*$ and increase the value of x_U until the corresponding $\hat{M} \approx \hat{B}$.

2.3 Context: Korean Residential Electricity Market

Although notches are useful to understanding consumer behaviors and preferences, research of notches in consumption goods markets is limited. The residential electricity market in Korea experienced major price reforms in 2016, which generated sizeable jumps in total prices. I use this variation, combined with rich data, to determine whether consumers respond to either average or marginal prices. I also use the notches to estimate not only price elasticity parameters, but optimization frictions.

2.3.1 Electricity Market in Korea

Korea's electricity market is regulated highly; the government owns the monopolistic electric utility Korea Electric Power Corporation (KEPCO). Although KEPCO is a publicly listed firm (Korea Exchange in 1989 and New York Stock Exchange in 1994), the South Korean government holds 51.1 % of KEPCO's share,[1] and the company's status is regulated by the Korea Electric Power Corporation Act.

KEPCO operates in both wholesale and retail markets, and it covers the entire country's electricity grid system. From generation through transmission to distribution, KEPCO directly or indirectly controls the supply and prices. KEPCO participates in the generation market through its subsidiaries, which generate the majority of electricity[2][3]. KEPCO purchases electricity from the generators under the supervision of the operator, Korea Power Exchange (KPX). KEPCO is thus the sole supplier in transmission, distribution, and retail markets.

Every consumer on the electricity grid is classified by KEPCO as residential, industrial, general, agricultural, educational, midnight power, or street lights. It retails to over 23 million customers, 15 million of which are residential (KEPCO 2020a). The residential sector outnumbers all others in terms of number of customers, but the industrial sector consumes the majority of energy provided.

2.3.2 Nonlinear Residential Electricity Prices

Each customer in disparate sectors experiences different price structures and contract options. To study consumption behavior under notches, I focus on residential customers due to the nonlinearity of the prices they pay and data availability.

Unlike with commercial electricity, residential tariffs are nonlinear[4] and use increasing-

1. More details about the structure of the company is available at KEPCO (2020b).

2. The subsidiaries generated over 70% of electricity in 2018. Gross energy produced was 593 Terawatt hours (TWh, equal to 10^12 watt hours), and KEPCO's subsidiaries were responsible for 418 TWh. The subsidiaries also hold the majority of generating capacity; they have a combined capacity of 81 Gigawatts (GW, equal to 10^9 watt), and the national total is 123 GW. For more details on gross generation and capacity, see KEPCO (2019).

3. Other participants in the generation market include governmental agencies, such as The Korea Water Resources Corporation, and private firms. (KEPCO 2019)

4. In my sample period, all residential tariffs are fixed rate as opposed to real-time or time-of-use pricing. Residential electricity prices are provided each month, and they do not vary across days or hours. General, industrial, and educational customers can opt for time-of-use rates and pay lower prices during off-peak

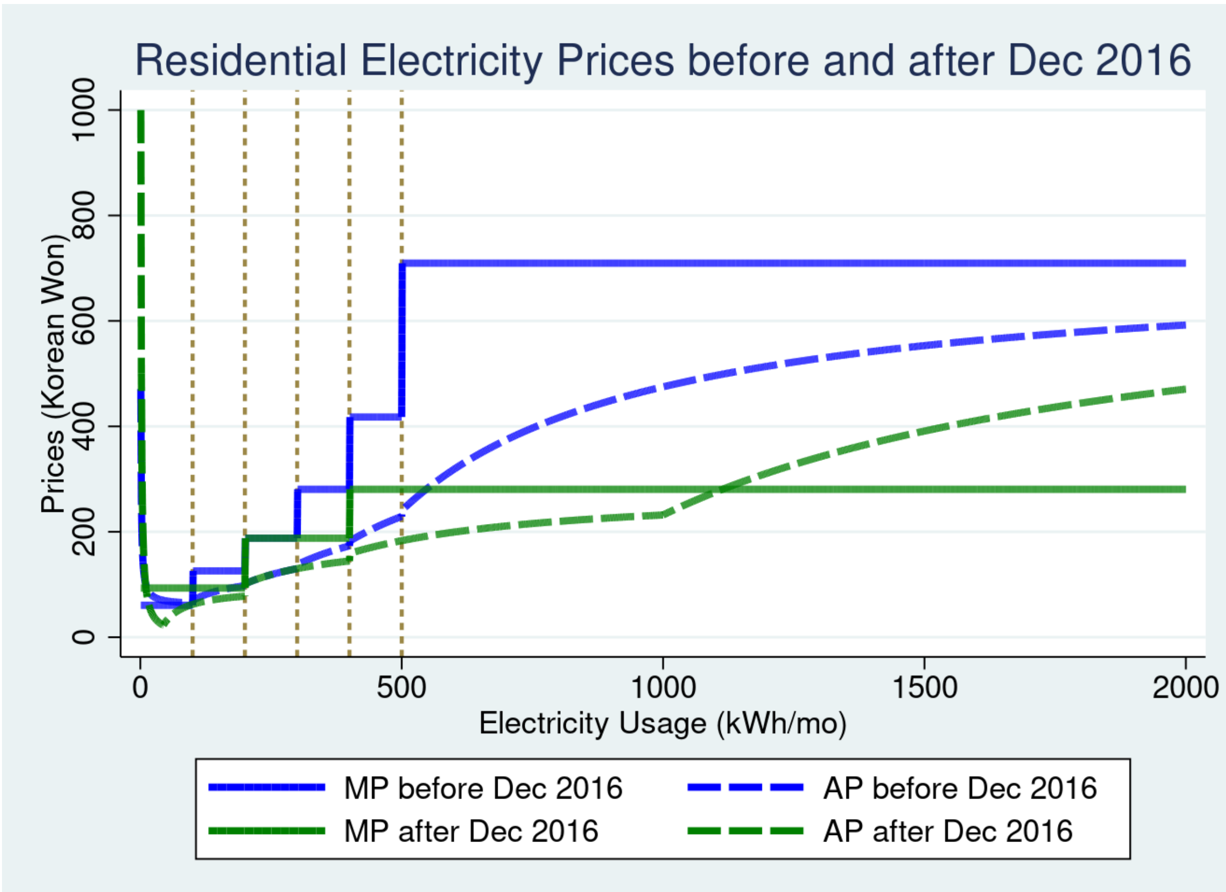


Figure 2.1: Residential Electricity Prices Before and After 2016

Note: This figure plots electricity prices, marginal and average, for low-voltage residential customers. The blue lines describe prices before the price reform in Dec 2016, and the green lines depict the prices after the reform. Solid lines represent the marginal prices for one kilowatt-hour in Korean Won. Dashed lines represent the average prices of the electricity the consumers have paid up to the monthly usage level. Throughout this paper, I assume 1 USD = 1100 KRW.
Source: Price data provided by KEPCO (2016).

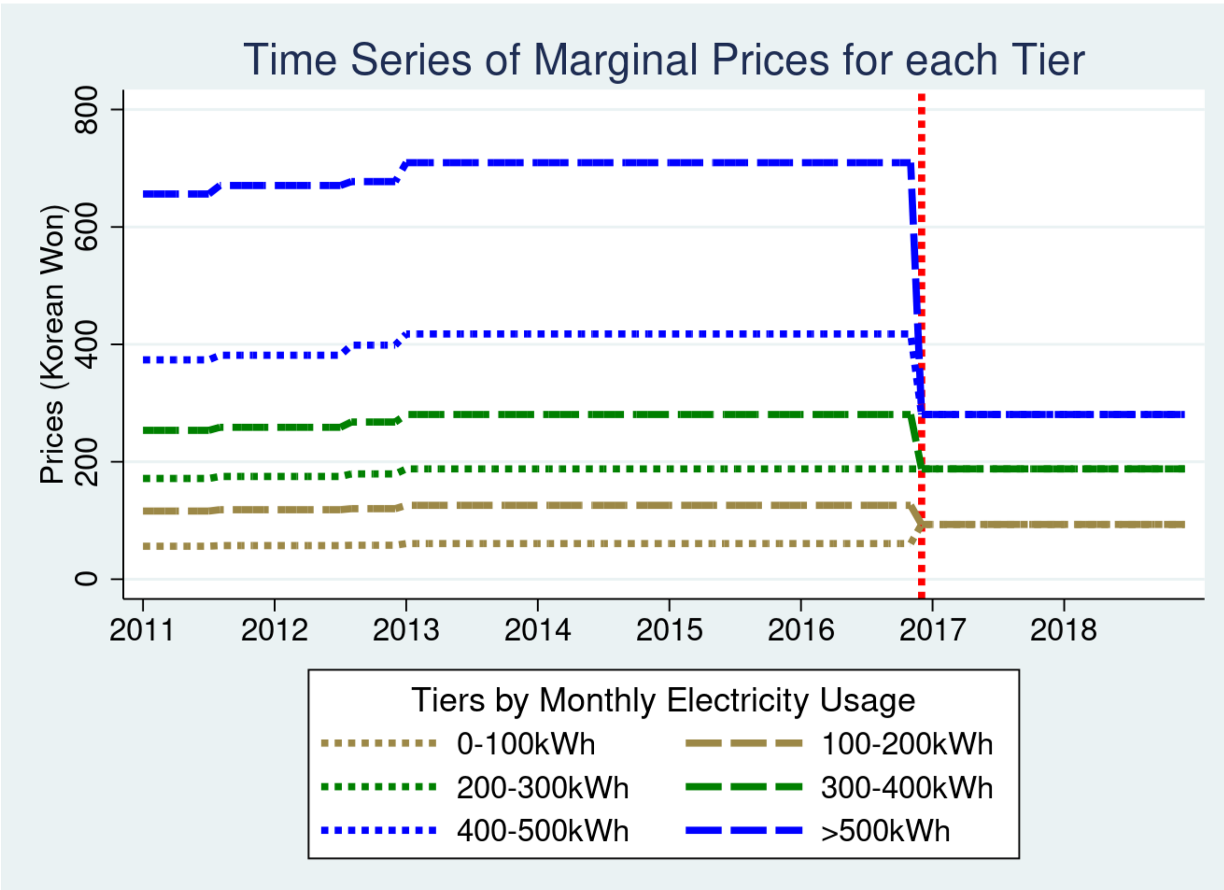


Figure 2.2: Time Series Variation of Marginal Prices

Note: This figure plots marginal electricity prices for residential low-voltage customers. For each tiers that face different marginal prices, I draw the time series of the marginal prices from Jan 2011 to Dec 2018. I mark the major price change that happened in Dec 2016 in red dashed vertical line. The tiers that each line represent are described in the legend. The 0-100kWh tier (in brown dotted line) faces the lowest marginal prices throughout all periods, and the 500kWh or more tier (in blue dashed line) the highest. All lines show the prices for one kilowatt-hour in Korean Won. Throughout this paper, I assume 1 USD = 1100 KRW.
Source: Price data provided by KEPCO (2016).

block pricing (IBP). The marginal price is an increasing step function of monthly electricity use, hence the name increasing-block pricing, and price jumps are steep. For example, during December 2018, an increase from 198kWh to 199kWh cost 93.3 extra Korean Won (KRW)[5], but the cost for the same 1kWh more than doubles to 187.9 KRW at 200kWh. The lowest and highest marginal prices differ by up to 11.7 times in the sample[6]. Such differences in marginal prices are rarely observed in extant research[7]. My analyses identify whether customers still fail to internalize the price scheme under severe nonlinearity.

Residential customers are on either low- or high-voltage contracts, with slightly different price tariffs. I focus on low-voltage customers, since I can observe household-level electricity consumption. Low-voltage electricity (220V) can be used directly with household appliances, but high-voltage customers must step the electricity down (22,900V) using their own transformers. Due to this restriction, the high-voltage option typically serves multi-family dwellings, and most single-family homes operate under low-voltage contracts. Although low- and high-voltage residential tariffs have similar structures, low-voltage prices are slightly higher since they include extra maintenance costs, including transforming electricity to lower voltages. Electricity consumption data are available at the building level, and I remove multi-family households to ensure that observed electricity use represents household-level consumption. Consumers in the resulting subsample thus experience residential low-voltage pricing.

In addition to cross-sectional price variation from increasing block pricing, a major price reform during 2016 generated a considerable amount of time-series variation. In the past, consumers were informed of price changes in advance, changes were proportional, and overall

---

hours.

5. In my sample period of 2011 to 2018, the annual average exchange rate was between 1 USD = 1054 and 1161 KRW. Throughout this paper, I assume 1 USD = 1100 KRW.

6. From Nov 2013 to Nov 2016, customers in the 0-100 kWh range experienced a marginal price of 60.7 KRW ($\approx$ 5.52 cents) per 1kWh. The marginal price increases to 709.5 KRW ($\approx$ 64.5 cents) per 1kWh if they used more than 500kWh per month.

7. For example, variation in the marginal price of electricity was less than three times in Ito (2014).

price structures remained the same. However, on 13 December 2016, KEPCO substantially changed the residential tariff without prior notice, reducing the number of tiers from six to three. The gap between the marginal prices also reduced from 11.7 times to around three times. The price changes were retroactive and were applied to bills starting from 1 December 2016. Since the electricity utility was regulated heavily, this price change was neither a market-driven adjustment or a response to monthly demand. Customers were unlikely to anticipate the timing of the announcement or the retroactive nature of the price change. Discussions on the price change had not been disclosed to the public until final approval by the Electricity Regulatory Commission (KOREC). I use this unexpected, plausibly exogenous shock to the tariff structure as variation that identifies consumer behavior under nonlinear pricing.

Not only did the residential electricity payment jump significantly at multiple points, the notches were implicit and consumers needed to review their contracts closely to understand their total payments. The residential electricity price is nonlinear in the sense that the marginal price increases over use and there are discrete jumps in total payments. Increases to marginal electricity prices are delineated in the price chart and are well-informed. However, the price table does not articulate how the changes in basic charges, or fixed charges, affect the total payment. (See Figure A.2.) Hence, customers may misunderstand the "fixed" monthly charges as a truely fixed cost and overlook the notch points.

Discounts or changes to "fixed" monthly charges generated notch points and complicated bill calculation. Monthly base rates increased at 200 and 400 kWh, generating jumps in total payments. A discount of up to 4,000 KRW (about USD 3.63) was introduced during December 2016 for those with electricity consumption fewer than 200kWh. The discount and changes to "fixed" monthly charges made total payments jump by 31.4 % for 200kWh and 10.3 % for 400kWh[8]. Although changes to "fixed" prices over consumption blocks and

8. The residential electricity bill was 15,570 KRW for 200kWh and 20,447 KRW for 201kWh (before tax), resulting in an extra 4,877 KRW charge for 1kWh used. The discount accounted for 4,000 KRW, changes to base payments for 690 KRW, and marginal price of 1kWh for the remaining 187 KRW. Similarly, total

the discount generated substantial notches, they are included only as a footnote in the price chart.

The aforementioned discount and a surcharge during peak months further complicated marginal price calculations. An addendum to the residential electricity contract stated that the total electricity bill should not be lower than 1,000 KRW after the 4,000 KRW discount for the 0-200kWh range. This article makes the total payment flat for regions 0-43kWh, and the marginal price becomes zero. Another addendum implied surcharges during the summer (i.e., July and August) and during winter (i.e., December, January, and February), which increased the marginal price from 280.6 KRW to 709.5 KRW per kWh if a consumer used over 1,000 kWh.

Although a tariff structure can be affected by changes to "fixed" rates, the discount, and the surcharge, consumers might be unaware of these details if they do not read the terms thoroughly. I define naive total payment and naive marginal price as the nonlinear tariff constructed using only marginal prices in the price chart. Naive prices ignore the extra nonlinearity caused by jumps to fixed prices, the discount under 200kWh, and the seasonal surcharge for heavy users. AsIto (2014) suggests, the least-informed group responds only to average prices, defined as the ratio of total payment and total electricity consumption. I refer to them as average-price maximizers. The other extreme group comprises those who internalized the nonlinear structures of the tariff fully, called marginal-pricemaximizers. I classify those who internalized only the price chart as naive marginal-price maximizers. In the following sections, I empirically test whether implicit terms in the electricity tariffs hindered consumers from internalizing the nonlinear price structure fully.

payments increased from 65,760 KRW to 72,560 KRW around 400kWh, and the base payment increased from 1,600 KRW to 7,300 KRW.

2.4 Descriptive Data and Bunching Histogram

2.4.1 Data Sources

I observe building-level electricity use in nationwide data from 2011 to 2018. To ensure that each observation represents only one household, I select single-family residential customers and construct their perceived prices using electricity price data.

I construct data on electricity use and building information. Electricity use data includes the address, municipality code, and monthly electricity use (in kWh) of each building. Data are published by the Ministry of Land, Infrastructure and Transport of Korea (MOLIT), which also provides data on the primary purpose of the building, number of households, number of floors, area (square meters), energy efficiency, and other architectural information. I merge these data using address as a matching variable.

I select observations that lie under the nonlinear residential tariff, ensuring that each row in the resulting subsample represents only one customer. I use building type information to rule out non-residential properties. Where building-level use equals household-level consumption in single-family homes, it is impossible to separate unit-level consumption in multi-family dwellings in the data. Hence, I focus on single-family homes, eliminating multi-family dwellings using information on building type and number of households. Single-family homes pay low-voltage residential price tariffs except in the uncommon case in which they demand a large quantity and have set up their own transformer. To further refine the sample, I use information on maximum capacity. Under low-voltage contracts, consumers cannot use more than 3kW of electricity at a time. In an extreme case in which they use this full capacity all the time, monthly use is limited to 2,160 kWh[9]. I remove observations whose monthly electricity use exceeds 2,000 kWh, which ultimately removes observations unlikely to be low-voltage customers.

Based on residential price data published by KEPCO (2016), I generate marginal, av-

9. 3 kWh/hour $\times$ 24 hours/day $\times$ 30 days = 2,160 kWh

Table 2.1: Summary Statistics for Residential Electricity Data

	mean	sd	p5	p25	p50	p75	p95
Electricity Usage (kWh/mo)	376	297	43	178	294	487	1000
Building Area (sq m)	89	75	17	43	74	104	218
Marginal Price (in KRW)	324	322	61	126	188	418	710
Average Price (in KRW)	181	124	66	94	129	218	463
Naive Marginal Price (in KRW)	308	227	61	126	188	418	710
Total Price (in KRW)	102351	154815	2797	16171	37359	104973	454394
Number of Unique Buildings	191278						
Number of Observations	5159468						

Notes: I report the summary statistics of the residential electricity customers using the building-level monthly electricity data and the building information data published by MOLIT. I report the mean, standard deviation, and percentiles (5, 25, 50, 75, and 95) in each column. The monthly electricity usage is in Kilowatt-hours unit. The building area refers to the total square meterage of the floorplan. Marginal, average, naive marginal, and total prices are in Korean Won (KRW). In my sample period, the exchange rate was bounded by 1243.6 and 1009.2 USD/KRW. I use an approximate exchange rate of 1 USD = 1,100 KRW.

erage, and naive marginal prices at each consumption level for each month. Full price information is shown in the Electric Rates Table provided by KEPCO, including marginal prices and fixed payments for each consumption block. Discounts for fewer than 200kWh and seasonal surcharges are added in an appendix to the price chart. Using this information, I calculate the total payment and marginal price for an additional 1kWh. I also construct alternative prices for consumers who internalize the price charts and ignore the appendix (i.e., naive marginal price) and those who compare only total payments to total electricity use (i.e., average price).

2.4.2 Summary Statistics and Bunching Diagram

I report summary statistics of electricity use, the size of the building, and prices in Table 2.1. In addition to means and standard errors, I report percentiles since consumption is skewed to the right, shown in Figure 2.3. The first row of the table indicates that the majority of electricity consumption use falls between 100 and 500 kWh, where marginal and total prices jump. Not only do the notch and kink points affect most customers, the nonlinear jumps are economically significant in comparison to the median total price of 37,359. Residential electricity bills jump from 15,570 KRW to 20,447 KRW at 200kWh and from 65,760 KRW to

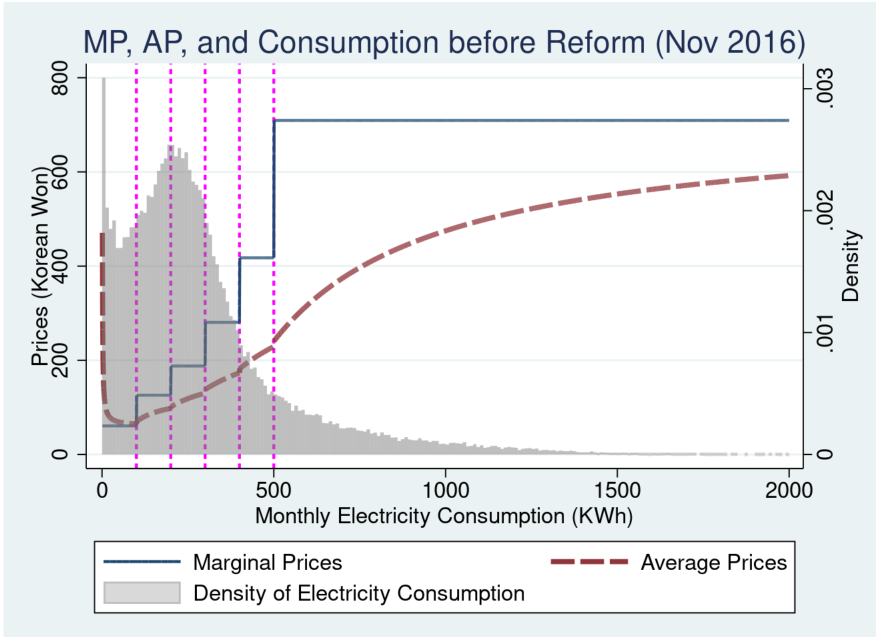
72,560 KRW at 400kWh[10]. Nonlinearity in the tariffs makes total prices even more positively skewed than electricity use. For example, the 75th percentile user consumes 2.74 times more electricity in comparison to the 25th percentile user (Row 1, Columns p5 and p75), but their payments differ by 6.49 times (row 6).

Consistent with extant literature (Borenstein 2009; Ito 2014), I do not observe strong bunching responses to nonlinear prices in Figure 2.3. The density of monthly electricity consumption is plotted as a gray histogram, and marginal and average prices at each consumption level (i.e., kWh per month) are depicted as solid and dashed lines, respectively. The residential electricity tariff previously consisted of six tiers with steep steps in the marginal price function (Panel a). The price reform effective 1 December 2016 reduced the number of tiers to three and substantially reduced disparities between the lowest and highest marginal prices (Panel b). Standard economic theory suggests that electricity consumption distribution bunches before nonlinear points (marked with dotted horizontal lines in magenta), given that consumers internalize marginal prices. However, no bunching is evident before points of nonlinearity. Even more antithetically, I observe excess distributions after the nonlinear point of 200kWh both before and after price reform. I therefore argue that consumers failed to respond to marginal prices and formally test whether they responded to alternative prices such as average prices or naive marginal prices.

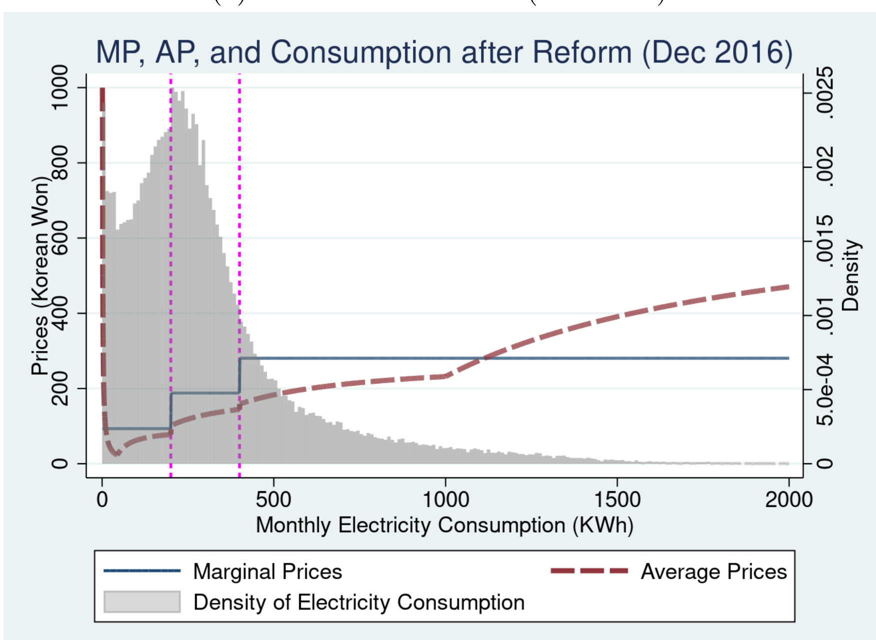
2.5 Encompassing Test of Perceived Prices

Preliminary analysis of the data and literature suggests that consumers do not internalize marginal prices fully but respond to alternative price measures. Using changes to the price tariff from December 2016 as an instrument, I find that consumers responded to average rather than marginal prices.

10. 1 USD = 1,100 KRW.



(a) Before Price Reform (Nov 2016)



(b) After Price Reform (Dec 2016)

Figure 2.3: Price Tariffs and Electricity Consumption Distribution

Note: This figure plots the distribution of residential customers along their monthly consumption and the marginal prices they face. Panel (a) shows the distribution and the marginal prices right before the price change in Dec 2016, and Panel (b) the afterwards. I truncate the distribution at 2000kWh per month, which is close to the absolute maximum a residential electricity customer with low-voltage contract could consume. The thresholds of each tier is marked with pink dotted vertical lines. I plot the marginal prices for each tier with blue solid lines, the average prices with red dashed lines. The price axes are on the left side. I plot the histogram of customers in grey bars, with the density axis on the right side. Throughout this paper, I assume 1 USD = 1100 KRW.
Source: Price data provided by KEPCO (2016).

2.5.1 Estimation Model

Recall that the quasi-linear, iso-elastic utility function in equation (2.1) leads to the first-order condition in equation (2.6). Taking logs, demand becomes:

$$logx_{it} = logn_i + e \times logp_{it}, \tag{2.20}$$

where x_it is monthly electricity use by individual i at time t, n_i is a preference parameter, and e is price elasticity of electricity demand. For rational consumer i who optimizes at the margin, p_it is the marginal price that i experiences at consumption level x_it. I allow consumers to respond to alternative price measures, and I let preference n for electricity consumption to vary by consumers by month m:

$$logx_{it} = logn_{i,m} + b_{MP} \times logMP_{it} + b_{AP} \times logAP_{it} \tag{2.21}$$

I also allow electricity use to grow over time and annual growth rate a_i to be heterogeneous across individuals. Accounting for potential time effects γ_t and the error term, I derive estimation equation:

$$logx_{it} = a_i \times t/12\gamma_t + logn_{i,m} + b_{MP} \times logMP_{it} + b_{AP} \times logAP_{it} + \epsilon_{it} \tag{2.22}$$

I take a 12-month difference to eliminate seasonal effects specific to each household:

$$\begin{aligned}
\Delta_{12}logx_{it} &:= logx_{it} - logx_{i,t-12} \\
&= a_i \times (t-(t-12))/12 + (\gamma_t - \gamma_{t-12}) + (logn_{i,m} - logn_{i,m}) \\
&+ b_{MP} \times (logMP_{it} - logMP_{i,t-12}) \\
&+ b_{AP} \times (logAP_{it} - logAP_{i,t-12}) + (\epsilon_{it} - \epsilon_{i,t-12}) \\
&= a_i + c_t + b_{MP} \times \Delta_{12}logMP_{i,t} + b_{AP} \times \Delta_{12}logAP_{i,t} + e_{it}
\end{aligned} \tag{2.23}$$

2.5.2 Identification Strategy

This model controls for all individual-by-season effects and allows an individual to use electricity more in general and/or to use electricity more during specific seasons. For example, the model accounts for the possibility that an individual uses more electricity for air conditioning than for heating in comparison to other individuals. The model also allows for linear changes in the electricity consumption distribution. Using this estimation model, I use an encompassing test (Davidson and MacKinnon 1981) to empirically determine to what price measure consumers respond. If, for example, all consumers respond only to average prices, b_{AP} will be different from zero and b_{MP} close to zero. In addition to marginal and average price pairs, I test combinations of marginal, average, and naive marginal price combinations and report results.

One challenge to estimation is intrinsic simultaneity among the dependent variable, demand, and the explanatory variable, prices. I solve this problem by isolating price changes due to the policy from the effects of demand on prices. Marginal, average, and naive marginal prices are, by construction, the tariff schedules evaluated at electricity use x_{it}:

$$MP_{it} = MP_t(x_{it}) \tag{2.24}$$

In my data, tariff schedule $MP_t(\cdot)$ is exogenous and changes disproportionately. The change itself is uncorrelated with error term e_{it} as long as argument $\tilde{x}$ is independent of e_{it}. If I take prior electricity use $\tilde{x}$ as the argument, the entire term $MP_t(\tilde{x})$ is uncorrelated with e_{it}. Public finance literature commonly constructs policy-induced instruments, or simulated instruments,[11] based on this term.

$$logMP_{it}^{PIIV} = logMP_t(\tilde{x}) - logMP_{t-12}(\tilde{x}) \tag{2.25}$$

11. Feldstein (1995) and Saez, Slemrod, and Giertz (2012) assess the influences of changes to tax tariff structures. Ito (2014) and Shaffer (2019) use this approach in electricity markets to test alternative prices.

Using this policy-induced instrument variable, I fix consumption level $\tilde{x}$ based on the prior consumption histories of each individual and isolate the influence of changes to price schedules on actual prices. I take a 12-month difference to eliminate seasonal effects, and take the midpoint of six month lagged electricity use $\tilde{x} = x_{t-6}$ as the argument. The instrument is exogenous since its components $logMP_t(\tilde{x})$ and $logMP_{t-12}(\tilde{x})$ are exogenous. This instrument is relevant since the jump in $MP_t(\cdot)$ affects both $MP_{it} = MP_t(x_{it})$ and $MP_t(\tilde{x})$. The null of weak instruments is rejected during the first stage regression of $logMP_{it}^{PIIV}$, $logAP_{it}^{PIIV}$, and $logNMP_{it}^{PIIV}$.

Another source of endogeneity occurs when consumption is non-stationary and grows heterogeneously. A clear, cross-sectional discontinuity in price tariffs means that a researcher can use one location as a control group and estimate distributional changes in another region. In my data, all consumers are under the same price schedule and thus this method is not viable. Alternatively, I allow the growth rate to be individual-specific but assume the heterogeneous growth rate is linear. Under this assumption, individual-specific effect a_i not only captures distributional changes fully but allows me to estimate the growth rate specific to each individual directly.

2.5.3 Estimation Results

I estimate main estimation equation (2.23) using various specifications, reporting results in Tables A.9-A.10. The IV regressions suggest that people respond strongly to average prices but the marginal and naive marginal price elasticities are close to zero or even positive.

Table 2.2 shows estimation results from the preferred specification—IV estimation with both individual and time fixed effects. Results account for endogeneity issues from simultaneity and non-stationary distribution of electricity use. The dependent variable is monthly electricity use in kWh, and the explanatory variables are marginal, average, and naive marginal prices, or their combinations. The coefficients indicate respective elasticities, since variables on both sides are in logs. I also report the standard errors of the estimators in parentheses.

Table 2.2: Estimated Price Elasticity of Demand: IV, with Time and Individual Fixed Effects

	(1) none/none	(2) none/none	(3) none/none	(4) none/none	(5) none/none	(6) none/none	(7) none/none
marginal price	0.115***			0.218***	0.186***		0.171***
	(0.010)			(0.018)	(0.028)		(0.030)
average price		-0.047**		-0.207***		-0.132***	-0.256***
		(0.022)		(0.025)		(0.031)	(0.040)
naive marginal price			0.043***		-0.068***	0.090***	0.069*
			(0.010)		(0.023)	(0.016)	(0.037)
Obs	3467682	3479456	3479456	3467682	3467682	3479456	3467682

Notes: This table reports the price elasticity of electricity demand. I report the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. The sample period is from Jan 2011 to Dec 2018. The data contains 191,278 unique addresses. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use marginal, average, and naive marginal prices as price measures, and each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors are clustered at address level and are shown in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01

I cluster errors at address level in order to account for serial correlation.

Columns 1-3 show that people internalize average prices but respond to marginal and naive marginal prices in the opposite direction of what economic theory predicts. The average price elasticity of demand is -4.7 percents, which is significantly different from zero. However, a one percent increase in marginal price leads to a 0.115 percent increase in consumption.

Results are more singular in the regressions of both average and marginal prices. In Columns 4-6, I regress electricity use with pairwise combinations of marginal, average, and naive marginal prices. Average prices further absorb the effect of marginal and naive marginal prices. Consumers tended to respond more to naive marginal prices in comparison to marginal prices (Column 5). Column 7 shows that only average prices decrease electricity use but marginal and naive marginal prices increase electricity use when I account for all price measures.

I also report results based on the model without individual and time fixed effects in Table A.9. Although the coefficients are not positive for marginal and naive marginal prices under this specification (Columns 1 and 3), they become positive once I account for average prices (Columns 4, 6, and 7). Average price elasticity estimates remain negative across the specifications. The standard errors are smaller in comparison to those in Table 2.2 since the individual fixed effects substantially reduce degrees of freedom. However, omission of the fixed effects might create bias, and I continue to include these fixed effects. Average price coefficient estimates are relatively robust to inclusion of fixed effects, but elasticity estimates from marginal and naive marginal prices might be a result of a spurious regression. For reference, I include ordinary least squares results in Tables A.11 and A.10. Unsurprisingly, estimates suffer severe upward bias due to reverse causality.

Table 2.3: Longer Run Responses to Prices, with Individual and Time Fixed Effects

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE
marginal price	0.066***		
	(0.015)		
lagged marginal price (t-1)	0.004		
	(0.020)		
lagged marginal price (t-2)	0.111***		
	(0.018)		
lagged marginal price (t-3)	-0.053		
	(0.042)		
lagged marginal price (t-4)	0.112***		
	(0.036)		
average price		-0.045	
		(0.028)	
lagged average price (t-1)		0.010	
		(0.034)	
lagged average price (t-2)		0.048	
		(0.031)	
lagged average price (t-3)		0.109	
		(0.093)	
lagged average price (t-4)		-0.075	
		(0.095)	
naive marginal price			-0.006
			(0.013)
lagged naive marginal price (t-1)			0.026*
			(0.014)
lagged naive marginal price (t-2)			0.153***
			(0.017)
lagged naive marginal price (t-3)			-0.225***
			(0.078)
lagged naive marginal price (t-4)			0.330***
			(0.073)
Obs	3088564	3098870	3098870

Notes: This table reports the longer-run lagged demand responses to electricity prices. I report the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. I use one- to four-month lagged marginal, average, and naive marginal prices as the explanatory variables. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. Due to the lags up to four months, the new sample period is from May 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(3) represents the estimation results based on the price measure specifications. Standard errors are clustered at address level and are shown in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01

Table 2.4: Responses to Lagged Prices, with Individual and Time Fixed Effects

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE	(4) Coef./SE	(5) Coef./SE	(6) Coef./SE	(7) Coef./SE
lagged marginal price (t-1)	0.128***			0.208***	0.100***		0.096***
	(0.011)			(0.020)	(0.032)		(0.034)
lagged average price (t-1)		-0.006		-0.147***		-0.130***	-0.250***
		(0.023)		(0.026)		(0.032)	(0.042)
lagged naive marginal price (t-1)			0.104***		0.026	0.151***	0.154***
			(0.010)		(0.026)	(0.016)	(0.041)
Obs	3375474	3383345	3383345	3375474	3375474	3383345	3375474

Notes: This table reports the lagged demand responses to electricity prices. I report the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. I use one-month lagged marginal, average, and naive marginal prices as the explanatory variables. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. Due to the one-month lag, the new sample period is from Feb 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors are clustered at address level and are shown in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01

2.5.4 Lagged Responses

If consumers respond to lagged prices, the current marginal and naive marginal prices might not have strong relationships with demand, and thus I investigate whether demand responses are lagged.

In Table 2.4, I estimate demand responses to one-month lagged prices. The signs of the lagged estimates are consistent with current price estimates. The average price estimate is no longer significant (Column 2), and results are similar without fixed effects (Table A.9).

Table 2.3 estimates coefficients for longer lags. Column 1 reports demand responses to current and past marginal prices with up to four months lags. No evidence suggests that previous marginal prices reduce current electricity use. Similarly, Columns 2 and 3 report results for lagged average and naive marginal prices. Over the lags, marginal prices have mixed signs regarding the influence on electricity use. Price elasticity with respect to current average prices, however, continues to be negative. The elasticity estimate of -0.045 is very similar to the estimate without the lagged terms (Column 2, Table 2.2). Any lags in average prices do not explain electricity demand.

2.6 Conclusions

Theory suggests that consumers are more likely to bunch around notches in comparison to kinks because the former are a stronger form of nonlinearity in prices in the sense that they contain kinks as the mildest case without jumps in total payment. This also suggests that consumers who do not perceive nonlinear prices with kinks fully might internalize a price schedule once notches are introduced.

Contrary to this prediction, this paper suggests that consumers still fail to internalize a full price structure and respond to an alternative price measure, average price, which is simpler to calculate. This paper thus serves as an example of studying the role of notch points in consumption goods markets. I also show that the price elasticity of demand with respect to average prices is relatively robust across specifications. An estimated elasticity of -0.047 accords with an estimate from Southern California of -0.051 (Ito 2014). However, the estimate is not a structural estimate and elasticity might differ in other markets and contexts.

Consumers respond to marginal prices unconventionally. In Figure 2.3 I observe bunching behavior at the right instead of left of the notch point. One explanation is that consumers notice that their use has exceeded the notch point and thus subsequently and regretfully reduce consumption.

From a policy perspective, the nonlinear price in my data is not socially efficient; although the marginal cost of electricity is sloping upward in real-time (Borenstein 2002), no evidence suggests that the marginal cost increases at the monthly level. Policymakers behind nonlinear electricity prices, or increasing block prices, have been blamed for creating inefficiency by confusing these two concepts. Such inefficiency is attenuated when consumers respond to average instead of marginal prices.

Another finding is that when some information on prices is difficult to understand, consumers partially internalize the price schedule. By comparing naive marginal and marginal prices, I test whether the way price structure details are presented matters. In my sample, consumers responded more to an easier version of marginal prices. Experimenting with partial and full information provision in nonlinear pricing contexts would provide more insights for rational inattention literature.

REFERENCES

Allcott, Hunt and Michael Greenstone (2017). *Measuring the welfare effects of residential energy efficiency programs*. Tech. rep. National Bureau of Economic Research.

Allcott, Hunt and Daniel Keniston (2017). “Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America”. In: *The Review of Economic Studies* 85.2, pp. 695–731.

Allcott, Hunt and Christopher Knittel (2019). “Are consumers poorly informed about fuel economy? Evidence from two experiments”. In: *American Economic Journal: Economic Policy* 11.1, pp. 1–37.

Allcott, Hunt and Dmitry Taubinsky (2015). “Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market”. In: *American Economic Review* 105.8, pp. 2501–38.

Bartik, Alexander W et al. (2019). “The local economic and welfare consequences of hydraulic fracturing”. In: *American Economic Journal: Applied Economics* 11.4, pp. 105–55.

Bartik, Timothy J (1991). *Who benefits from state and local economic development policies?* WE Upjohn Institute for Employment Research.

Blanchard, Olivier Jean et al. (1992). “Regional evolutions”. In: *Brookings papers on economic activity* 1992.1, pp. 1–75.

Bolsen, Toby and James N Druckman (2018). “Do partisanship and politicization undermine the impact of a scientific consensus message about climate change?” In: *Group Processes & Intergroup Relations* 21.3, pp. 389–402.

Bolsen, Toby, James N Druckman, and Fay Lomax Cook (2014). “The influence of partisan motivated reasoning on public opinion”. In: *Political Behavior* 36.2, pp. 235–262.

Borenstein, Severin (2002). “The trouble with electricity markets: understanding California’s restructuring disaster”. In: *Journal of economic perspectives* 16.1, pp. 191–211.

— (2009). “To what electricity price do consumers respond? Residential demand elasticity under increasing-block pricing”. In: *Preliminary Draft April* 30, p. 95.

Borick, Christopher P and Barry G Rabe (2017). "Personal experience, extreme weather events, and perceptions of climate change". In: *Oxford Research Encyclopedia of Climate Science*.

Boudet, Hilary et al. (2016). "The effect of industry activities on public support for 'fracking'". In: *Environmental Politics* 25.4, pp. 593–612.

Brunner, Eric, Stephen L Ross, and Ebonya Washington (2011). "Economics and policy preferences: causal evidence of the impact of economic conditions on support for redistribution and other ballot proposals". In: *Review of Economics and Statistics* 93.3, pp. 888–906.

Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer (2013). "Are consumers myopic? Evidence from new and used car purchases". In: *American Economic Review* 103.1, pp. 220–56.

Carvalho, Leandro and Dan Silverman (2019). *Complexity and Sophistication*. Tech. rep. National Bureau of Economic Research.

Census Bureau – Census Bureau, U.S. Department of Commerce (2019a). *2019 TIGER/Line Shapefiles (machine-readable data files)*.

Census Bureau – — (2019b). *2019 TIGER/Line Shapefiles Technical Documentation*.

Cook, John et al. (2016). "Consensus on consensus: a synthesis of consensus estimates on human-caused global warming". In: *Environmental Research Letters* 11.4, p. 048002.

Cook, John et al. (2013). "Quantifying the consensus on anthropogenic global warming in the scientific literature". In: *Environmental research letters* 8.2, p. 024024.

Cragg, Michael I et al. (2013). "Carbon geography: the political economy of congressional support for legislation intended to mitigate greenhouse gas production". In: *Economic Inquiry* 51.2, pp. 1640–1650.

Davidson, Russell and James G MacKinnon (1981). "Several tests for model specification in the presence of alternative hypotheses". In: *Econometrica: Journal of the Econometric Society*, pp. 781–793.

Demski, Christina et al. (2017). “Experience of extreme weather affects climate change mitigation and adaptation responses”. In: *Climatic Change* 140.2, pp. 149–164.

Deryugina, Tatyana and Olga Shurchkov (2016). “The effect of information provision on public consensus about climate change”. In: *PloS one* 11.4, e0151469.

Druckman, James N, Erik Peterson, and Rune Slothuus (2013). “How elite partisan polarization affects public opinion formation”. In: *American Political Science Review* 107.1, pp. 57–79.

EIA – Energy Information Administration, U.S. Department of Energy (2011). *Review of emerging resources: US Shale gas and shale oil plays.*

EIA – — (2016). *Maps: Exploration, Resources, Reserves, and Production.* URL: `https://www.eia.gov/maps/maps.htm`. (accessed: Oct.10.2019.)

EIA – — (2019a). *Annual energy outlook 2019.*

EIA – — (2019b). *Energy Explained: Where our natural gas comes from.* URL: `https://www.eia.gov/energyexplained/natural-gas/where-our-natural-gas-comes-from.php`. (accessed: Oct.10.2019.)

Fedaseyeu, Viktar, Erik Gilje, and Philip E Strahan (2015). *Voter preferences and political change: Evidence from shale booms.* Tech. rep. National Bureau of Economic Research.

Feldstein, Martin (1995). “The effect of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act”. In: *Journal of Political Economy* 103.3, pp. 551–572.

Feyrer, James, Erin T Mansur, and Bruce Sacerdote (2017). “Geographic dispersion of economic shocks: Evidence from the fracking revolution”. In: *American Economic Review* 107.4, pp. 1313–34.

Gifford, Robert (2011). “The dragons of inaction: Psychological barriers that limit climate change mitigation and adaptation.” In: *American psychologist* 66.4, p. 290.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2018). *Bartik instruments: What, when, why, and how.* Tech. rep. National Bureau of Economic Research.

Hausman, Jerry A (1979). “Individual discount rates and the purchase and utilization of energy-using durables”. In: *The Bell Journal of Economics*, pp. 33–54.

Howe, Peter D et al. (2014). “Mapping the shadow of experience of extreme weather events”. In: *Climatic change* 127.2, pp. 381–389.

Howe, Peter D et al. (2015). “Geographic variation in opinions on climate change at state and local scales in the USA”. In: *Nature Climate Change* 5.6, p. 596. URL: `https://climatecommunication.yale.edu/visualizations-data/ycom-us/`.

Ito, Koichiro (2014). “Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing”. In: *American Economic Review* 104.2, pp. 537–63.

Kahn, Matthew E and Matthew J Kotchen (2011). “Business cycle effects on concern about climate change: the chilling effect of recession”. In: *Climate Change Economics* 2.03, pp. 257–273.

Kahn, Matthew E and Frank A Wolak (2013). *Using information to improve the effectiveness of nonlinear pricing: Evidence from a field experiment*. California Air Resources Board, Research Division.

Kearney, Melissa S and Riley Wilson (2018). “Male earnings, marriageable men, and nonmarital fertility: Evidence from the fracking boom”. In: *Review of Economics and Statistics* 100.4, pp. 678–690.

KEPCO (2016). *Electricity Price*. URL: `http://home.kepco.co.kr/kepco/EN/main.do`. (accessed: Mar.20.2020).

— (2019). *Statistics of Electric Power in Korea*. URL: `https://home.kepco.co.kr/`.

— (2020a). *KEPCO Big Data Center*. URL: `https://bigdata.kepco.co.kr`. (accessed: Mar.20.2020).

— (2020b). *KEPCO Overview*. URL: `https://home.kepco.co.kr/`. (accessed: Mar.20.2020).

Kleven, Henrik J and Mazhar Waseem (2013). “Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan”. In: *The Quarterly Journal of Economics* 128.2, pp. 669–723.

Leeper, Thomas J and Rune Slothuus (2014). "Political parties, motivated reasoning, and public opinion formation". In: *Political Psychology* 35, pp. 129–156.

Liebman, Jeffrey B and Richard J Zeckhauser (2004). "Schmeduling". In:

McRae, Shaun and Robyn Meeks (2016). "Price perception and electricity demand with nonlinear tariffs". In: *mimeo*.

Mildenberger, Matto and Anthony Leiserowitz (2017). "Public opinion on climate change: Is there an economy–environment tradeoff?" In: *Environmental Politics* 26.5, pp. 801–824.

Mukherjee, Deep and Mohammad Arshad Rahman (2016). "To drill or not to drill? An econometric analysis of US public opinion". In: *Energy policy* 91, pp. 341–351.

Olson-Hazboun, Shawn K, Peter D Howe, and Anthony Leiserowitz (2018). "The influence of extractive activities on public support for renewable energy policy". In: *Energy policy* 123, pp. 117–126.

Rosenberg, Stacy et al. (2010). "Climate change: a profile of US climate scientists' perspectives". In: *Climatic Change* 101.3-4, pp. 311–329.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz (2012). "The elasticity of taxable income with respect to marginal tax rates: A critical review". In: *Journal of economic literature* 50.1, pp. 3–50.

Sallee, James M (2014). "Rational inattention and energy efficiency". In: *The Journal of Law and Economics* 57.3, pp. 781–820.

Shaffer, Blake (2019). "Misunderstanding Nonlinear Prices: Evidence from a Natural Experiment on Residential Electricity Demand". In: *American Economic Journal: Economic Policy*.

NORC – Smith, Tom W., Davern, Michael, Freese, Jeremy, and Morgan, Stephen L. (2019). *General Social Surveys, 1972-2018 [machine-readable data file]*, 3, 758 pp.

Tvinnereim, Endre and Elisabeth Ivarsflaten (2016). "Fossil fuels, employment, and support for climate policies". In: *Energy Policy* 96, pp. 364–371.

APPENDIX A

GSS SURVEY QUESTIONS

A.1 GSS Surveys on Environmental Belief

Question. “Improving and protecting environment (natenvir)”

Questions associated with this variable:

We are faced with many problems in this country, none of which can be solved easily or inexpensively. I’m going to name some of these problems, and for each one I’d like you to name some of these problems, and for each one I’d like you to tell me whether you think we’re spending too much money on it, too little money, or about the right amount. First (READ ITEM A) . . . are we spending too much, too little, or about the right amount on (ITEM)?

B. Improving and protecting the environment

Answers. 1: Too little 2. About right 3. Too much 8. Don’t know 9. No answer 0. Not applicable

Question. “whether they think environmental scientists agree on global warming (sciagrgw)”

Questions associated with this variable:

Please look at Card B3. On a scale of 1 to 5, where 1 means “Near complete agreement” and 5 means “No agreement at all”, to what extent do environmental scientists agree among themselves about the existence and causes of global warming?

Answers. 1: Near complete agreement 2-4. (between 1 and 5, larger meaning skepticism of agreement) 5. No agreement at all 8. Don’t know 9. No answer 0. Not applicable

Question. "whether they think environmental scientists well understand global warming (gwsci)"

Questions associated with this variable:

The first issue is global warming. Global warming means a trend toward warmer temperatures throughout the world, with more extreme weather in many places and changes in food production that could affect our way of life. Some people believe that the burning of gasoline and other fossil fuels causes global warming. Others say that global warming has purely natural causes. Please look at Card B2. On a scale of 1 to 5, where 1 means "Very Well" and 5 means "Not at All", how well do the following groups understand the causes of global warming?

A. Environmental Scientists

Answers. 1: Very well 2-4. (between 1 and 5, larger meaning skepticism of scientists) 5. Not at all 8. Don't know 9. No answer 0. Not applicable

A.2 GSS Surveys on Elections

Question. "who did you vote for in 2008 presidential election(PRES08)"

Questions associated with this variable:

K. In 2008, you remember that Obama ran for President on the Democratic ticket against McCain for the Republicans. Do you remember for sure whether or not you voted in that election?

1. IF VOTED: Did you vote for Obama or McCain?

Answer: 1. Obama 2. McCain 3. Other candidate 4. Didn't vote 8. Don't know 9. No answer 0. Not applicable

Question. "who you would have voted for in 2008 presidential election (if08who)"

Questions associated with this variable:

K. In 2008, you remember that Obama ran for President on the Democratic ticket against McCain for the Republicans. Do you remember for sure whether or not you voted in that election?

2. IF DID NOT VOTE OR INELIGIBLE:

Who would you have voted for, for President, if you had voted? Answer: 1. Obama 2. McCain 3. Other candidate 8. Don't know 9. No answer 0. Not applicable

Question. ”who did you vote for in 2004 presidential election(PRES04)”

Questions associated with this variable:

J. In 2004, you remember that Kerry ran for President on the Democratic ticket against Bush for the Republicans. Do you remember for sure whether or not you voted in that election?

1. IF VOTED: Did you vote for Kerry or Bush?

Answer: 1. Kerry 2. Bush 3. Nader 6. Didn't vote 8. Don't know 9. No answer 0. Not applicable

Question. “who you would have voted for in 2004 presidential election (if04who)”

Questions associated with this variable:

J. In 2004, you remember that Kerry ran for President on the Democratic ticket against Bush for the Republicans. Do you remember for sure whether or not you voted in that election?

2. IF DID NOT VOTE OR INELIGIBLE: Who would you have voted for, for President, if you had voted?

Answer: 1. Kerry 2. Bush 3. Nader 8. Don't know or remember 9. No answer 0. Not applicable

A.3 GSS Surveys on Occupation

The responses are in categorical variables with four-digit industry codes. The details on the codes and labels are available in the GSS codebook (NORC 2019) or more easily at https://gssdataexplorer.norc.org/variables/17/vshow.

Question. "Respondent's Occupation (indus10)"

Questions associated with this variable:

Respondent's Occupation, Prestige, and Industry (NAICS 2007 code)

Question. "Spouse's industry code (spind10)"

Questions associated with this variable:

Respondent's spouse's Occupation, Prestige, and Industry (NAICS 2007 code)

Question. "Father's industry code (paind10)"

Questions associated with this variable:

Respondent's father's Occupation, Prestige, and Industry (NAICS 2007 code)

Question. "Mother's industry code (paind10)"

Questions associated with this variable:

Respondent's mother's Occupation, Prestige, and Industry (NAICS 2007 code)

APPENDIX B

ADDITIONAL FIGURES

Foreign Residents   About Bill   **Electric Rates Calculation**

Electric Rates Calculation

Residential Service

The Electric rate may be calculated based on the electric rate table and formula

The amount above includes the Electric Power Industry Foundation Fund and surcharges excluding the television reception fee, penalty, and other fees.

Amount Consumed ______ kWh   Date ______ ~ ______

Rates Calculation

Electric Rates Table

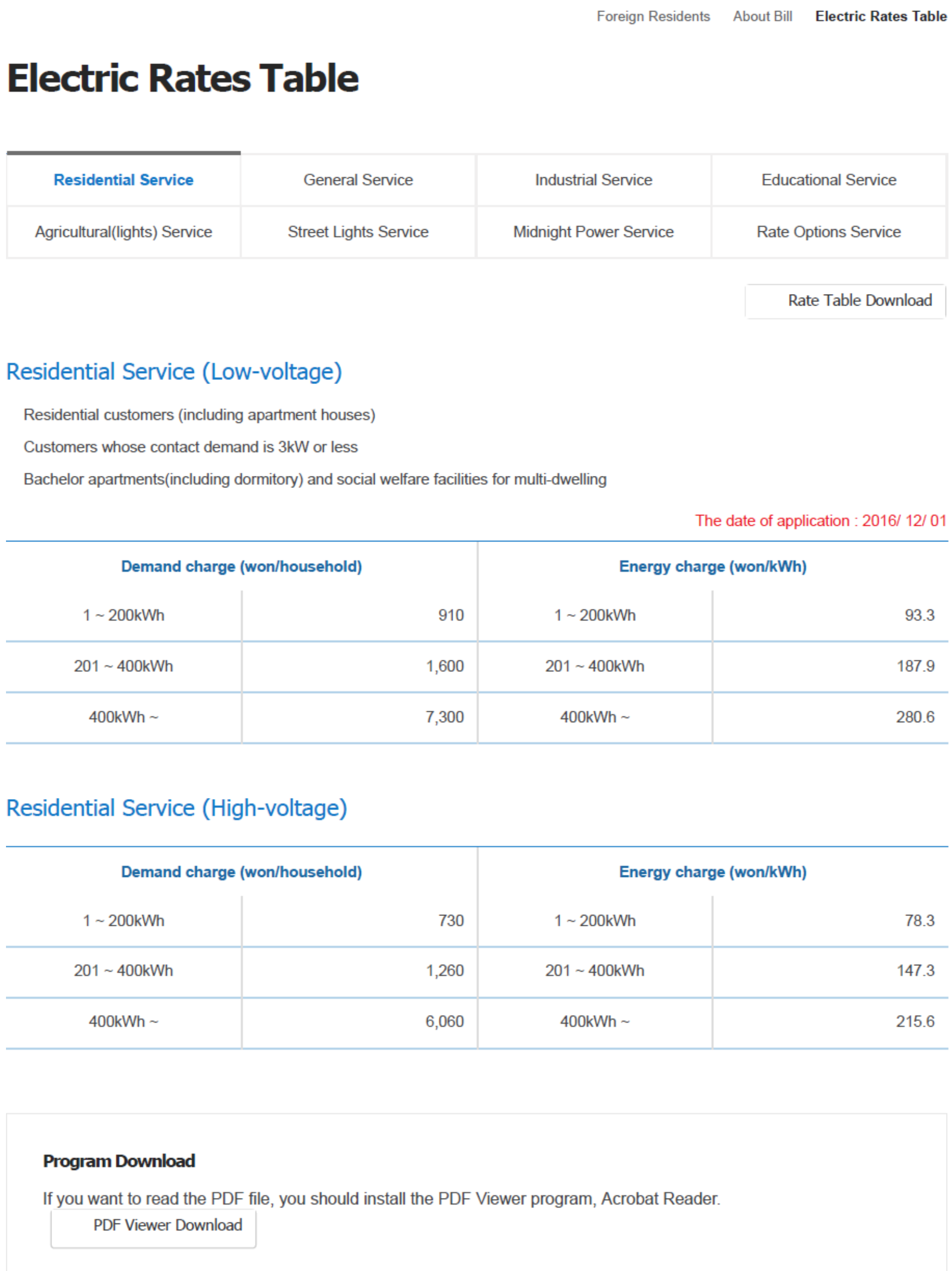
Basic charge (per premise)		Power demand charge (per kWh)	
For 200 kWh or less	730	For the first 200 kWh	78.3
For 201~400 kWh	1,260	For the next 200 kWh	147.3
For all over 400 kWh	6,060	For all over 400 kWh	215.6

Electric Rates Table by the Amount Consumed

Amount consumed		
1 ~ 300kWh	301 ~ 600kWh	601 ~ 1,000kWh

Figure A.1: Residential Electricity Price Calculator

Note: This is a screenshot of the electric rates calculation section of KEPCO website. This calculator offers consumers a way to compute their electricity prices. This website is public and can be accessed by anyone without signing in. If a user enters a monthly electricity level in the "Amount Consumed" item, set the period of the year in the "Date" item, and click the "Rates Calculation" button, the rates are calculated. Throughout this paper, I assume 1 USD = 1100 KRW.
Source: Price data provided by KEPCO (2016).
`http://home.kepco.co.kr/kepco/EN/F/B/ENFBPP002.do?menuCd=EN060202` Accessed: May 01, 2020



Foreign Residents About Bill **Electric Rates Table**

Electric Rates Table

Residential Service	General Service	Industrial Service	Educational Service
Agricultural(lights) Service	Street Lights Service	Midnight Power Service	Rate Options Service

Rate Table Download

Residential Service (Low-voltage)

Residential customers (including apartment houses)

Customers whose contact demand is 3kW or less

Bachelor apartments(including dormitory) and social welfare facilities for multi-dwelling

The date of application : 2016/ 12/ 01

Demand charge (won/household)		Energy charge (won/kWh)	
1 ~ 200kWh	910	1 ~ 200kWh	93.3
201 ~ 400kWh	1,600	201 ~ 400kWh	187.9
400kWh ~	7,300	400kWh ~	280.6

Residential Service (High-voltage)

Demand charge (won/household)		Energy charge (won/kWh)	
1 ~ 200kWh	730	1 ~ 200kWh	78.3
201 ~ 400kWh	1,260	201 ~ 400kWh	147.3
400kWh ~	6,060	400kWh ~	215.6

Program Download

If you want to read the PDF file, you should install the PDF Viewer program, Acrobat Reader.

PDF Viewer Download

Figure A.2: Residential Electricity Online Price Chart

Note: This is a screenshot of the electric rates table section of KEPCO website. This chart offers consumers another way to compute their electricity prices, on top of the rate calculator in Figure A.1. This website is public and can be accessed by anyone without signing in. A user can view the formula to calculate the electricity bill for each contract type. Throughout this paper, I assume 1 USD = 1100 KRW.
Source: Price data provided by KEPCO (2016).
`http://home.kepco.co.kr/kepco/EN/F/htmlView/ENFBHP00101.do?menuCd=EN060201` Accessed: May 01, 2020

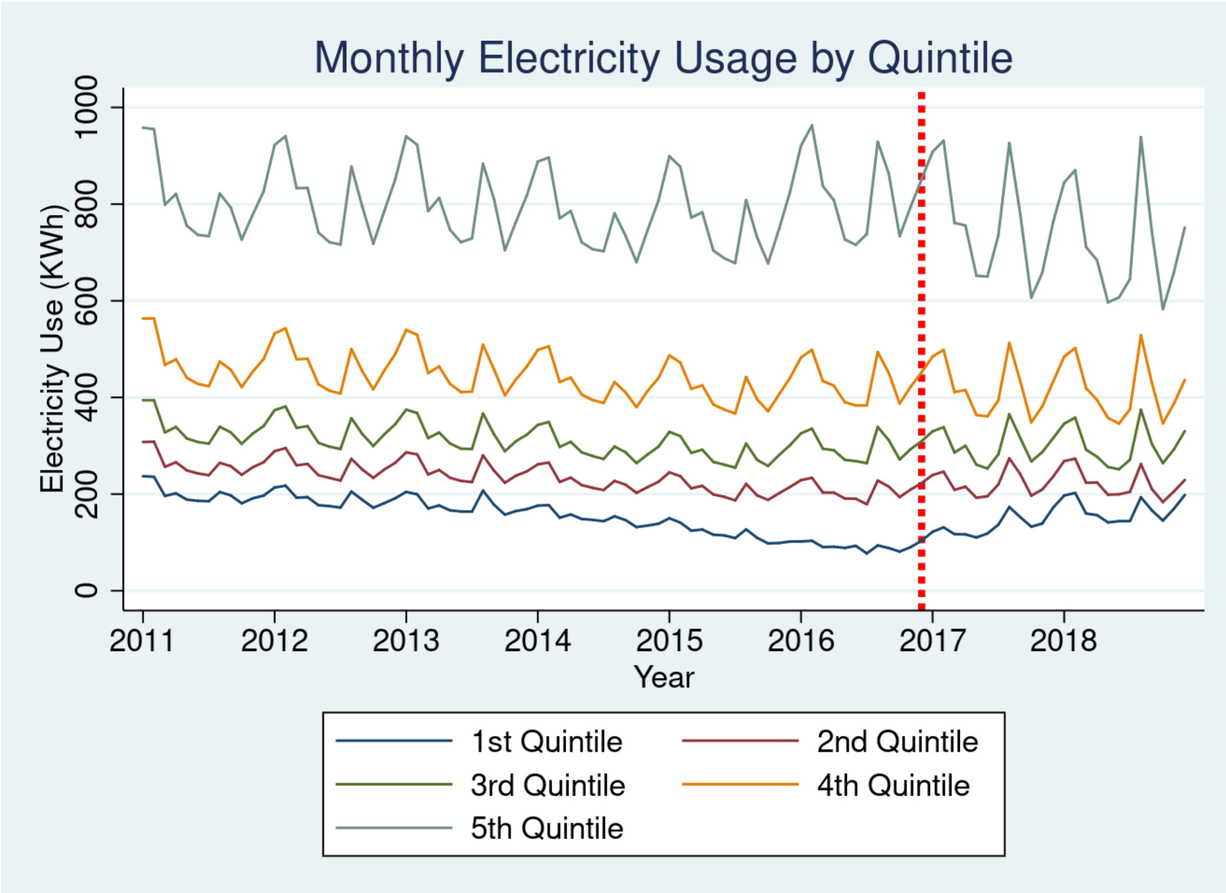


Figure A.3: Monthly Electricity Usage, by quintile

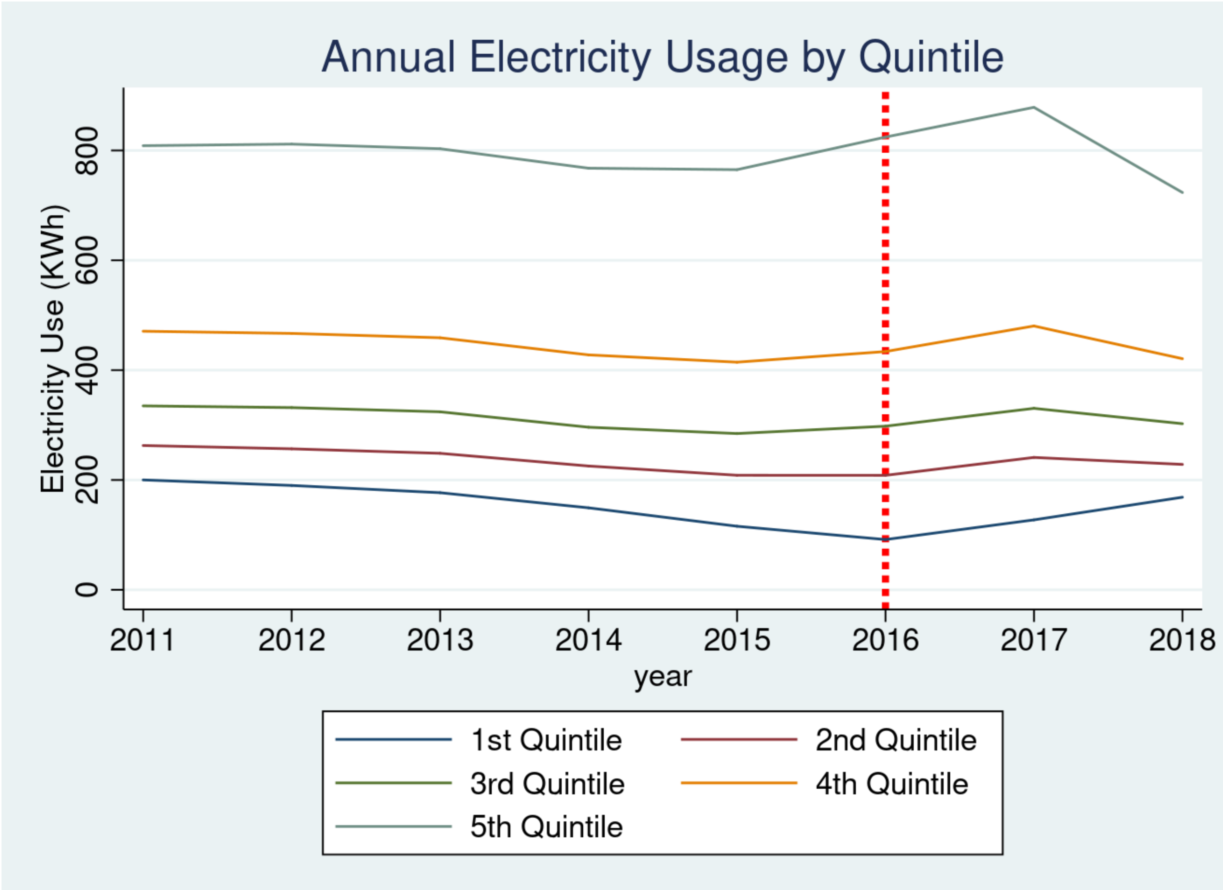


Figure A.4: Annual Electricity Usage, by quintile

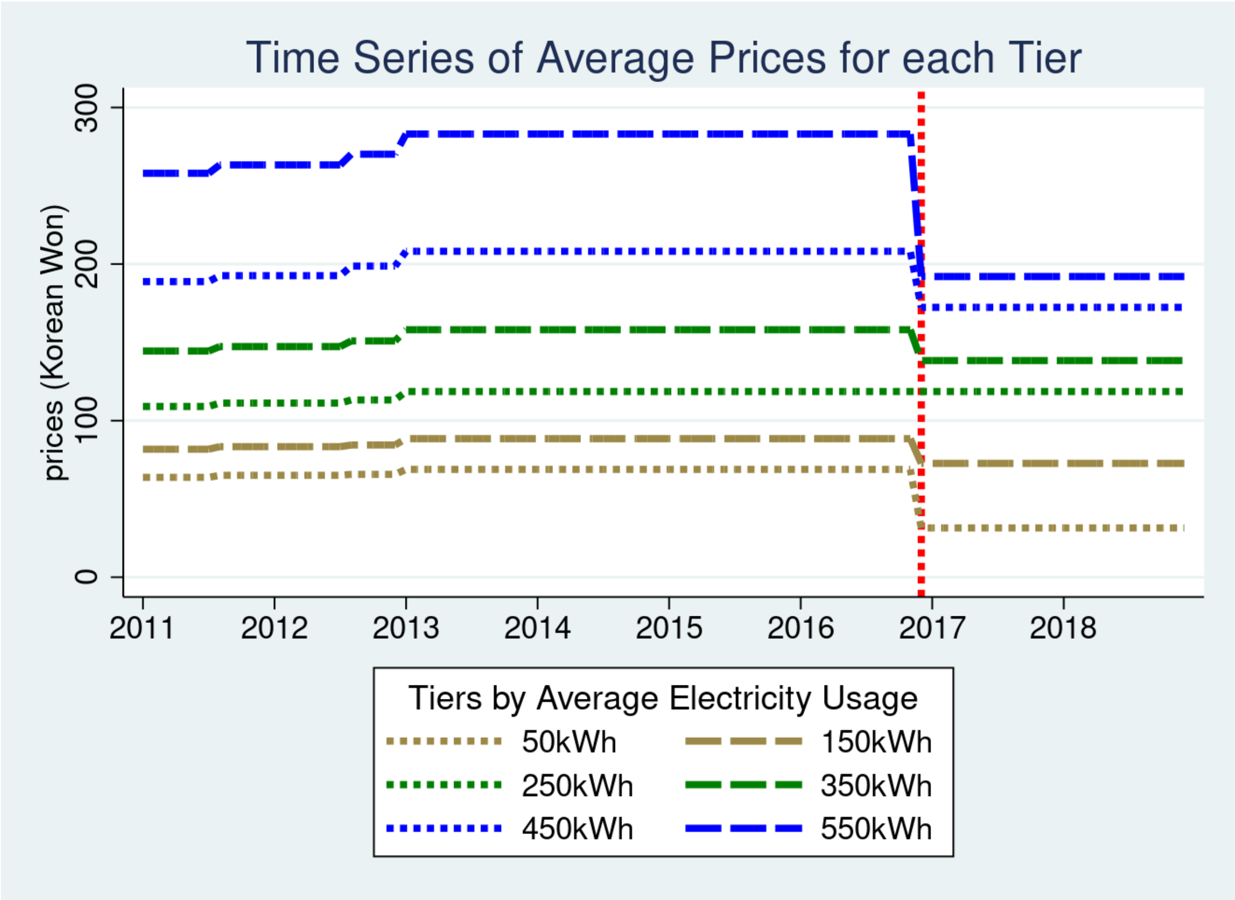


Figure A.5: Time Series of Average Prices, by tier

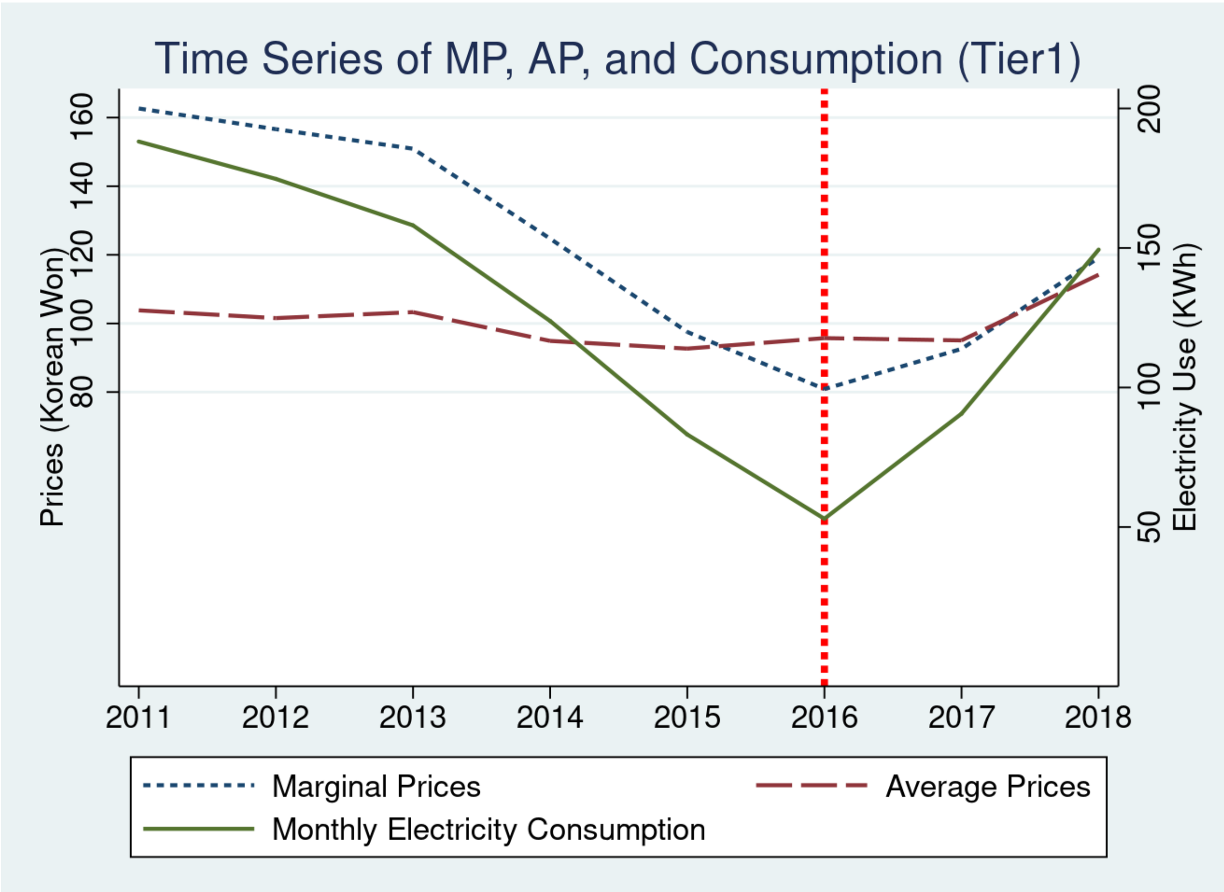


Figure A.6: Prices and Electricity Usage, Tier 1

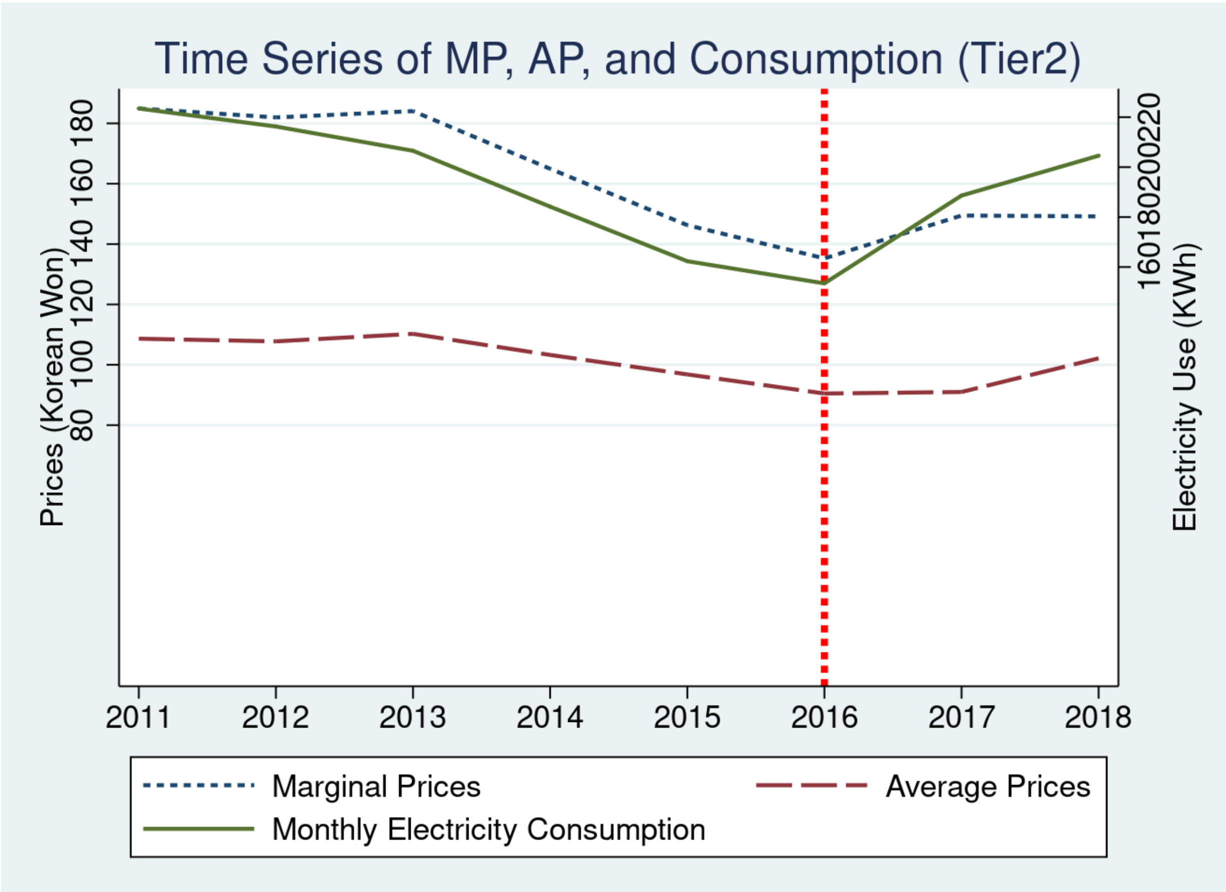


Figure A.7: Prices and Electricity Usage, Tier 2

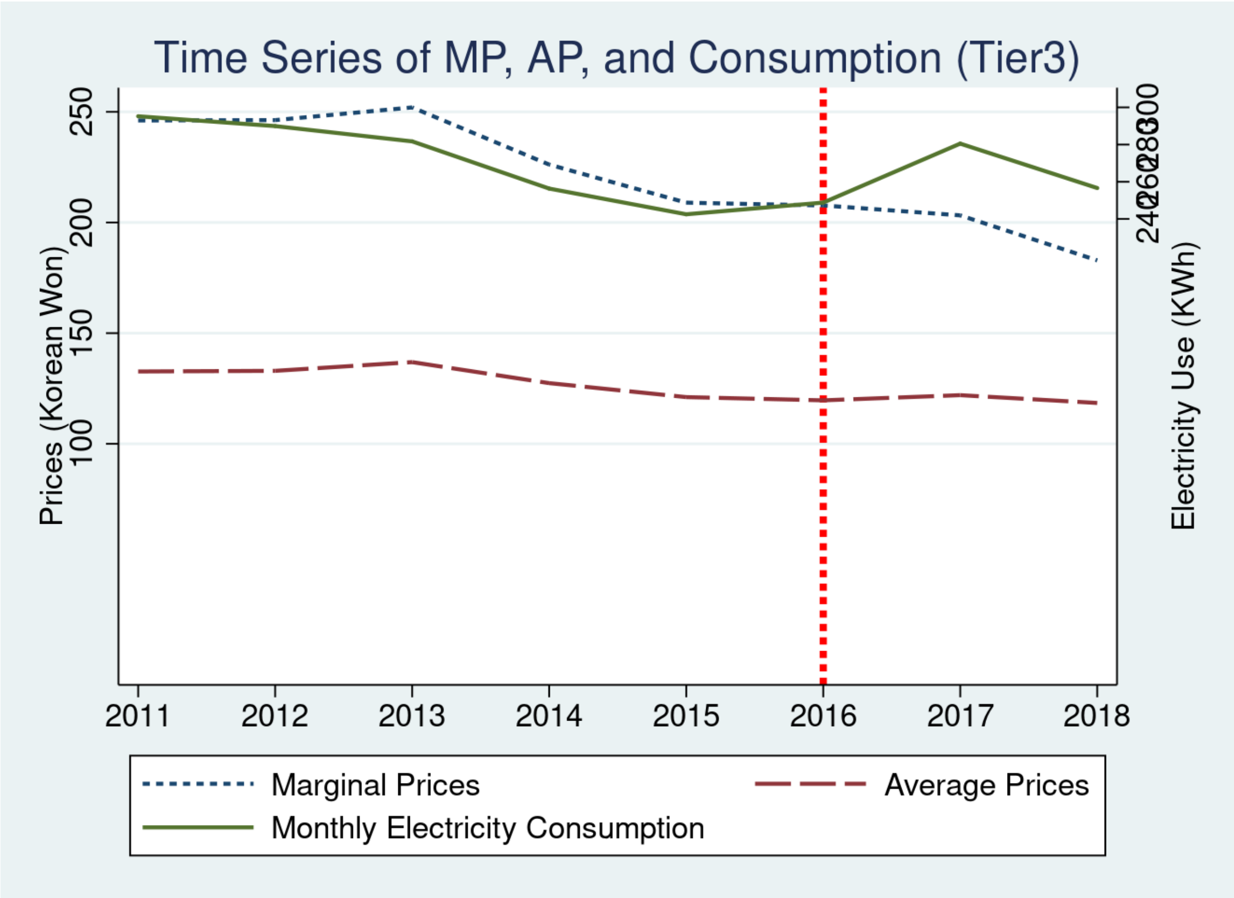


Figure A.8: Prices and Electricity Usage, Tier 3

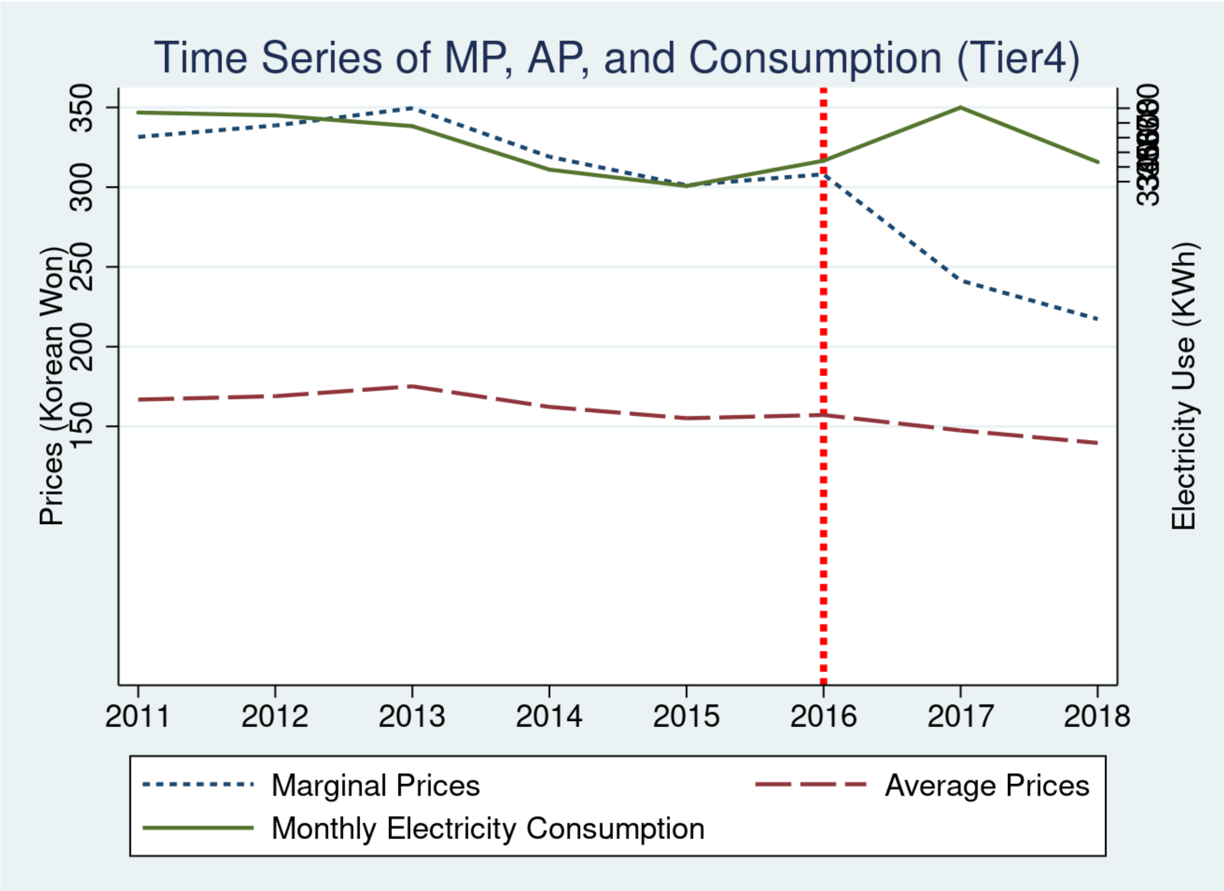


Figure A.9: Prices and Electricity Usage, Tier 4

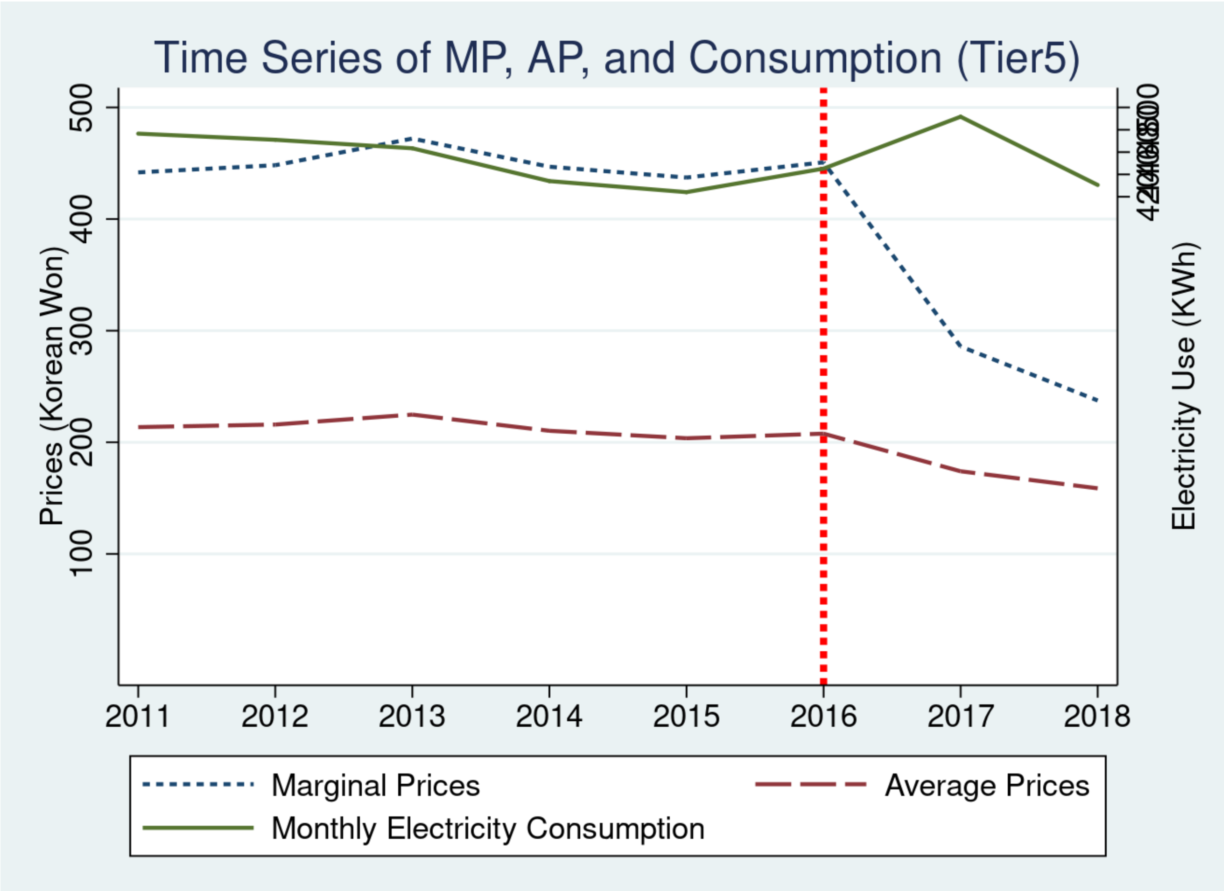


Figure A.10: Prices and Electricity Usage, Tier 5

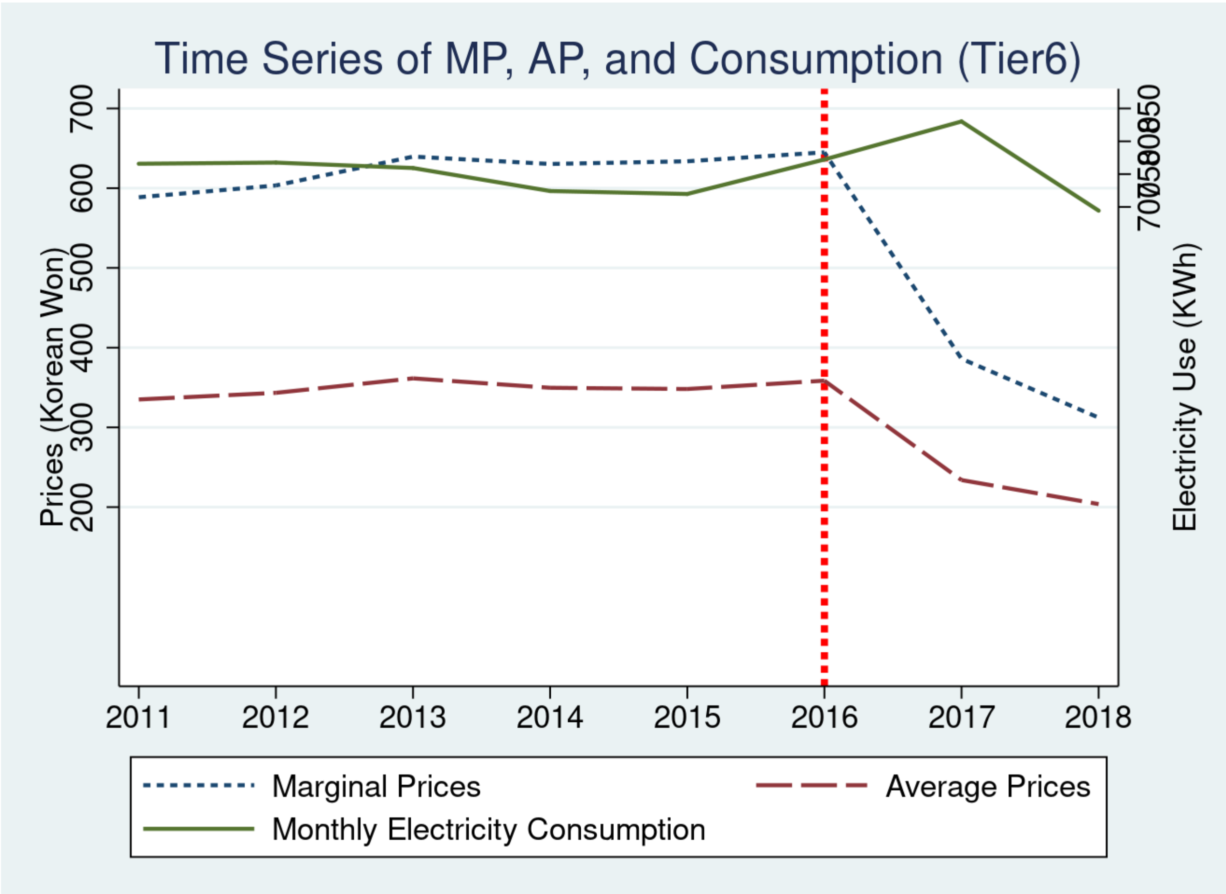


Figure A.11: Prices and Electricity Usage, Tier 6

APPENDIX C

ADDITIONAL TABLES

Table A.1: Generation and Capacity by Year

Years	2011	2012	2013	2014	2015	2016	2017	2018
Panel A. Generation Capacity (in MW)								
KEPCO and Subsidiaries	67,006	68,848	70,845	72,305	73,282	79,217	82,132	81,362
Others	12,336	12,957	16,124	20,911	24,367	26,649	34,775	37,729
Independent Power Producer	3,921	4,044	4,109	3,709	3,941	3,924	3,941	4,005
Total Capacity	83,263	85,849	91,077	96,925	101,590	109,789	120,848	123,096
Panel B. Gross Yearly Generation (in GWh)								
KEPCO and Subsidiaries	443,409	448,516	448,757	442,914	432,758	436,314	426,484	418,327
Others	52,671	58,964	66,710	78,495	94,757	103,906	126,742	151,520
Independent Power Producer	21,490	23,722	21,702	18,970	20,287	20,764	23,187	23,559
Total Generation	517,570	531,202	537,169	540,379	547,802	560,985	576,412	593,407

Note: This table illustrates the market power of KEPCO in wholesale electricity market. I report the capacity in megawatts in Panel A and the gross annual generation amount in gigawatt-hours in Panel B. In the first rows of each Panel, I report the annual average value of the capacity and the annual gross generation by KEPCO and its subsidiaries. The second rows, "Others", indicate the generators that are on the grid and participate in the centralized trading system. The Independent Power Producers are the firms that generate primarily for their own use and do not intend to participate in the wholesale electricity market, although it is possible for them to trade excess electricity generated. The last rows indicate the total amount of all three types of generators. Throughout the sample period of 2011-2018, KEPCO and its subsidiaries take the vast majority of the capacity and the generation amount. Source: KEPCO (2019)

Table A.2: Frequency Table of Building Purpose

	Purpose of Building frequency	percentage	cumulative percentage
single-family homes	3685725	63.03	63.03
med-sized businesses	501828	8.58	71.62
small businesses	463447	7.93	79.54
multi-family dwellings	296647	5.07	84.62
storages and warehouses	262951	4.50	89.11
agriculture and vet	208152	3.56	92.67
factories	158442	2.71	95.38
others	61909	1.06	96.44
nursery and elderly care	35636	0.61	97.05
lodging	31885	0.55	97.60
education	28056	0.48	98.08
office facilities	25764	0.44	98.52
religious	24288	0.42	98.93
automobile	19962	0.34	99.27
hazardous material facilities	18976	0.32	99.60
entertainment and public assembly	6472	0.11	99.71
garbage disposal facilities	6432	0.11	99.82
retail	5626	0.10	99.92
recreational	4960	0.08	100.00
Total	5847158	100.00	

Notes: Using the data of all buildings in Korea, published by MOLIT, I describe the characteristics of each building. This data includes residential, commercial, and other types of buildings. I report the purpose of each building in this frequency table. This table is in a descending order in terms of relative frequency. Each building is registered with its purpose, and the raw data reports a five digit code to indicate the type (or the purpose) of the building. I report the 17 most common types and combine the remaining types as others. I decode the building purpose and translate it to English. I report the count in the first column, the relative frequency in the second column, and the cumulative frequency in the third column. For additional information, contact the author (mekey@uchicago.edu).

Table A.3: Frequency Table of Number of Families

	Number of Families frequency	percentage	cumulative percentage
0	2237009	38.26	38.26
1	2967432	50.75	89.01
2	154399	2.64	91.65
3	123523	2.11	93.76
4	87877	1.50	95.26
5	72989	1.25	96.51
6	38396	0.66	97.17
7	27363	0.47	97.64
8	24432	0.42	98.05
9	15601	0.27	98.32
≥ 10	98137	1.68	100.00
Total	5847158	100.00	

Notes: Using the data of all buildings in Korea, published by MOLIT, I describe the characteristics of each building. This data includes residential, commercial, and other types of buildings. I report the number of families living in each building in this frequency table. Family refers to a group of people that are related by blood, marriage, or adoption. I truncate the variable and report the values of 10 or more as one bin at the second last row. This table is in an ascending order in terms of the value of the variable. I report the count in the first column, the relative frequency in the second column, and the cumulative frequency in the third column.

Table A.4: Frequency Table of Number of Households

	Number of Households frequency	percentage	cumulative percentage
0	5289864	90.47	90.47
1	217855	3.73	94.19
2	25084	0.43	94.62
3	22258	0.38	95.00
4	19077	0.33	95.33
5	12233	0.21	95.54
6	28550	0.49	96.03
7	16257	0.28	96.31
8	79113	1.35	97.66
9	16726	0.29	97.95
≥ 10	120141	2.05	100.00
Total	5847158	100.00	

Notes: Using the data of all buildings in Korea, published by MOLIT, I describe the characteristics of each building. This data includes residential, commercial, and other types of buildings. I report the number of households living in each building in this frequency table. Household is a group of people that share each unit or building and they do not need to be a family. I truncate the variable and report the values of 10 or more as one bin at the second last row. This table is in an ascending order in terms of the value of the variable. I report the count in the first column, the relative frequency in the second column, and the cumulative frequency in the third column.

Table A.5: Frequency Table of Number of Stories

	Number of Stories frequency	percentage	cumulative percentage
0	40252	0.69	0.69
1	3469846	59.34	60.03
2	1211469	20.72	80.75
3	505873	8.65	89.40
4	377188	6.45	95.85
5	145934	2.50	98.35
6	36976	0.63	98.98
7	16562	0.28	99.26
8	9208	0.16	99.42
9	6040	0.10	99.52
$\geq$ 10	27810	0.48	100.00
Total	5847158	100.00	

Notes: Using the data of all buildings in Korea, published by MOLIT, I describe the characteristics of each building. This data includes residential, commercial, and other types of buildings. I report the number of stories in each building in this frequency table. Number of stories indicate the number of floors above the ground in each building. I truncate the variable and report the values of 10 or more as one bin at the second last row. This table is in an ascending order in terms of the value of the variable. I report the count in the first column, the relative frequency in the second column, and the cumulative frequency in the third column.

Table A.6: Frequency Table of Number of Units

	Number of units frequency	percentage	cumulative percentage
0	5734792	98.08	98.08
1	50353	0.86	98.94
2	10548	0.18	99.12
3	5175	0.09	99.21
4	4921	0.08	99.29
5	2578	0.04	99.34
6	2900	0.05	99.39
7	1872	0.03	99.42
8	2687	0.05	99.46
9	1683	0.03	99.49
$\geq$ 10	29649	0.51	100.00
Total	5847158	100.00	

Notes: Using the data of all buildings in Korea, published by MOLIT, I describe the characteristics of each building. This data includes residential, commercial, and other types of buildings. I report the number of units in each building in this frequency table. Number of units indicate the number of separate and enclosed spaces (dwelling, commercial, industrial, and others) within each building. I truncate the variable and report the values of 10 or more as one bin at the second last row. This table is in an ascending order in terms of the value of the variable. I report the count in the first column, the relative frequency in the second column, and the cumulative frequency in the third column.

Table A.7: First Stage Regression: MP, AP, and MVP

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE
PIIV MP	0.217***	-0.036***	-0.070***
	(0.006)	(0.003)	(0.005)
PIIV AP	-0.142***	0.501***	0.034***
	(0.009)	(0.005)	(0.008)
PIIV MVP	0.409***	0.192***	0.767***
	(0.007)	(0.004)	(0.006)
First Stage F	7974.73	40978.479	48897.678
Obs	3467682	3467682	3467682

Notes: I test for weak instruments in this table. The null hypothesis is that the instruments do not explain the variation in the regressors, and high F-statistics lead to the rejection of the null of weak IV. I report the coefficients and the F-statistics estimates from the first stage of the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. The dependent variables in this table are marginal, average, and naive marginal prices, which are the regressors in the second stage estimation. I regress the price measures on the policy-induced instruments described in equation (2.25). The regressors and the instruments are in 12-month log changes, as described in the main equation. I use panel data of addresses (households) and month. The sample period is from Jan 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(3) represents the regression of marginal, average, and naive marginal prices. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.8: Estimated Price Elasticity of Demand: IV, with Time and Individual Fixed Effects

	(1) none/none	(2) none/none	(3) none/none	(4) none/none	(5) none/none	(6) none/none	(7) none/none
marginal price	0.115***			0.218***	0.186***		0.171***
	(0.007)			(0.010)	(0.020)		(0.020)
average price		-0.047***		-0.207***		-0.132***	-0.256***
		(0.010)		(0.014)		(0.013)	(0.020)
naive marginal price			0.043***		-0.068***	0.090***	0.069***
			(0.005)		(0.015)	(0.006)	(0.022)
Obs	3467682	3479456	3479456	3467682	3467682	3479456	3467682

Notes: This table reports the price elasticity of electricity demand. I report the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. The sample period is from Jan 2011 to Dec 2018. The data contains 191,278 unique addresses. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use marginal, average, and naive marginal prices as price measures, and each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.9: Estimated Price Elasticity of Demand: IV, with no Fixed Effects

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE	(4) Coef./SE	(5) Coef./SE	(6) Coef./SE	(7) Coef./SE
marginal price	-0.017***			0.239***	0.103***		0.148***
	(0.004)			(0.011)	(0.021)		(0.021)
average price		-0.043***		-0.401***		-0.103***	-0.467***
		(0.005)		(0.014)		(0.010)	(0.015)
naive marginal price			-0.004		-0.109***	0.051***	0.121***
			(0.003)		(0.018)	(0.006)	(0.020)
Constant	0.005***	0.003***	0.003***	0.004***	0.005***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs	3469730	3481444	3481444	3469730	3469730	3481444	3469730

Notes: I report the IV regression results in the main estimation equation (2.23) without fixed effects. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. The sample period is from Jan 2011 to Dec 2018. The data contains 191,278 unique addresses. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use marginal, average, and naive marginal prices as price measures, and each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.10: Estimated Price Elasticity of Demand: OLS, with Time and Individual Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
marginal price	0.612***			0.401***		-0.192***	-0.189***
	(0.000)			(0.001)		(0.001)	(0.001)
average price		0.972***		0.502***	-0.242***		-0.245***
		(0.001)		(0.001)	(0.001)		(0.001)
naive marginal price			0.894***		1.030***	1.090***	1.227***
			(0.000)		(0.001)	(0.001)	(0.001)
Constant	0.006***	0.006***	0.005***	0.007***	0.005***	0.008***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs	3491126	3501665	3501665	3491126	3501665	3491126	3491126

Notes: I report the OLS regression results in the main estimation equation (2.23) with individual and time fixed effects. I use panel data of addresses (households) and month. The sample period is from Jan 2011 to Dec 2018. The data contains 191,278 unique addresses. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use marginal, average, and naive marginal prices as price measures, and each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.11: Estimated Price Elasticity of Demand: OLS, with no Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
marginal price	0.659***			0.482***	0.482***	-0.201***	-0.193***
	(0.000)			(0.001)	(0.001)	(0.001)	(0.001)
average price		0.970***		0.396***	0.396***		-0.380***
		(0.001)		(0.001)	(0.001)		(0.001)
naive marginal price			0.926***			1.124***	1.333***
			(0.000)			(0.001)	(0.001)
Constant	0.006***	0.007***	0.006***	0.007***	0.007***	0.009***	0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs	3493330	3503848	3503848	3493330	3493330	3493330	3493330

Notes: I report the OLS regression results in the main estimation equation (2.23) without fixed effects. I use panel data of addresses (households) and month. The sample period is from Jan 2011 to Dec 2018. The data contains 191,278 unique addresses. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use marginal, average, and naive marginal prices as price measures, and each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.12: Responses to Lagged Prices, with Individual and Time Fixed Effects

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE	(4) Coef./SE	(5) Coef./SE	(6) Coef./SE	(7) Coef./SE
lagged marginal price (t-1)	0.128***			0.208***	0.100***		0.096***
	(0.007)			(0.011)	(0.025)		(0.026)
lagged average price (t-1)		-0.006		-0.147***		-0.130***	-0.250***
		(0.011)		(0.015)		(0.014)	(0.023)
lagged naive marginal price (t-1)			0.104***		0.026	0.151***	0.154***
			(0.005)		(0.020)	(0.007)	(0.027)
Obs	3375474	3383345	3383345	3375474	3375474	3383345	3375474

Notes: This table reports the lagged demand responses to electricity prices. I report the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. I use one-month lagged marginal, average, and naive marginal prices as the explanatory variables. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. Due to the one-month lag, the new sample period is from Feb 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.13: Responses to Lagged Prices

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE	(4) Coef./SE	(5) Coef./SE	(6) Coef./SE	(7) Coef./SE
lagged marginal price (t-1)	0.019***			0.254***	0.012		0.074***
	(0.004)			(0.012)	(0.027)		(0.027)
lagged average price (t-1)		-0.015***		-0.356***		-0.154***	-0.458***
		(0.006)		(0.015)		(0.010)	(0.016)
lagged naive marginal price (t-1)			0.038***		0.006	0.122***	0.224***
			(0.004)		(0.023)	(0.007)	(0.025)
Constant	-0.005***	-0.006***	-0.006***	-0.006***	-0.005***	-0.006***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Obs	3377673	3385490	3385490	3377673	3377673	3385490	3377673

Notes: This table reports the lagged demand responses to electricity prices. I report the IV regression results in the main estimation equation (2.23) without fixed effects. I use one-month lagged marginal, average, and naive marginal prices as the explanatory variables. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. Due to the one-month lag, the new sample period is from Feb 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(7) represents the estimation results based on the combination of price measures. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.14: Longer Run Responses to Prices, with Individual and Time Fixed Effects

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE
marginal price	0.066***		
	(0.018)		
lagged marginal price (t-1)	0.004		
	(0.025)		
lagged marginal price (t-2)	0.111***		
	(0.020)		
lagged marginal price (t-3)	-0.053		
	(0.052)		
lagged marginal price (t-4)	0.112**		
	(0.051)		
average price		-0.045**	
		(0.020)	
lagged average price (t-1)		0.010	
		(0.029)	
lagged average price (t-2)		0.048*	
		(0.025)	
lagged average price (t-3)		0.109	
		(0.079)	
lagged average price (t-4)		-0.075	
		(0.077)	
naive marginal price			-0.006
			(0.011)
lagged naive marginal price (t-1)			0.026*
			(0.015)
lagged naive marginal price (t-2)			0.153***
			(0.013)
lagged naive marginal price (t-3)			-0.225***
			(0.053)
lagged naive marginal price (t-4)			0.330***
			(0.052)
Obs	3088564	3098870	3098870

Notes: This table reports the longer-run lagged demand responses to electricity prices. I report the IV regression results in the main estimation equation (2.23) with individual and time fixed effects. I use one- to four-month lagged marginal, average, and naive marginal prices as the explanatory variables. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. Due to the lags up to four months, the new sample period is from May 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(3) represents the estimation results based on the price measure specifications. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.

Table A.15: Longer Run Responses to Prices, no Fixed Effects

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE
marginal price	-0.039***		
	(0.010)		
lagged marginal price (t-1)	0.002		
	(0.014)		
lagged marginal price (t-2)	0.078***		
	(0.012)		
lagged marginal price (t-3)	0.031		
	(0.023)		
lagged marginal price (t-4)	-0.059***		
	(0.022)		
average price		-0.008	
		(0.010)	
lagged average price (t-1)		-0.022	
		(0.015)	
lagged average price (t-2)		0.041***	
		(0.013)	
lagged average price (t-3)		0.162***	
		(0.026)	
lagged average price (t-4)		-0.154***	
		(0.024)	
naive marginal price			-0.077***
			(0.007)
lagged naive marginal price (t-1)			0.034***
			(0.010)
lagged naive marginal price (t-2)			0.127***
			(0.009)
lagged naive marginal price (t-3)			-0.003
			(0.021)
lagged naive marginal price (t-4)			0.023
			(0.020)
Constant	-0.015***	-0.016***	-0.017***
	(0.000)	(0.000)	(0.000)
Obs	3090071	3100322	3100322

Notes: This table reports the longer-run lagged demand responses to electricity prices. I report the IV regression results in the main estimation equation (2.23) without fixed effects. I use one- to four-month lagged marginal, average, and naive marginal prices as the explanatory variables. The dependent variable is monthly electricity consumption, and the explanatory variables are price measures. The dependent and explanatory variables are in 12-month log changes, as described in the main equation. I use the policy-induced instruments described in equation (2.25). I use panel data of addresses (households) and month. Due to the lags up to four months, the new sample period is from May 2011 to Dec 2018. The data contains 191,278 unique addresses. Each column in (1)-(3) represents the estimation results based on the price measure specifications. Standard errors in parentheses. Asterisks refer to significance levels; *: p-value <0.10, **: p-value <0.05, ***: p-value <0.01.