Title
Topics in Energy Economics, Environmental Economics, and Labor Economics

Author
Cui, Mengshan

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Topics in Energy Economics, Environmental Economics, and Labor Economics

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

by

Mengshan Cui

2019

ABSTRACT OF THE DISSERTATION

Topics in Energy Economics, Environmental Economics,
and Labor Economics

by

Mengshan Cui
Doctor of Philosophy in Economics
University of California, Los Angeles, 2019
Professor Dora Luisa Costa, Chair

This thesis consists of three chapters and covers topics in Energy Economics, Environmental Economics, and Labor Economics. The first chapter estimates the demand and supply functions of rooftop solar panels in California using a rebate program called the California Solar Initiative (CSI). Accurately estimating demand and supply is crucial for evaluating previous incentive programs and guiding future ones. I estimate the demand elasticity in California to be –3.284 and the supply elasticity to be 5.572. My contribution to the literature is a new method of using rebates as a source of exogeneous change to estimate both demand and supply functions simultaneously. I analyze disaggregated data at the Zipcode-month level, and I use a two-part model to incorporate large amounts of Zipcode-months with no solar panel installations. The second chapter is a joint work with Professor Edward Leamer and Jonathan Gu. We study the impact of Pasadena minimum wage on earnings, employment, and number of establishment. We use data from the individual zipcodes within and around Pasadena to conduct analysis. We find evidence of a positive impact of California/Pasadena minimum wages on the earnings of restaurant workers and of other low wage industries. Our model implies that a minimum wage increase of 10% would increase the average quarterly earnings per worker in limited-service restaurants by 8% and in full-service restaurants by 5%. Impact of minimum wage on employment and number of firms are less pronounced. The third chapter studies the impact of a Chinese environmental policy: Two-Control Zone(TCZ). I answer three questions in this study: One, has the TCZ policy been effective in reducing industrial emissions? Two, how has the TCZ policy affected economic activities? And three, how has the TCZ policy affected industry composition? To unpack these issues, I investigate whether a city designated as TCZ improved its environmental performance, city-level GDP growth, and level of foreign direct investment; had higher industrial output; and saw changes in industry composition. Using propensity-score matching to solve the possible endogeneity problem, I find evidence that TCZ cities had 5% lower ambient pollution from 2000 to 2012. During the same period, the TCZ policy caused 8% lower GDP growth and almost 3% less new foreign direct investment in TCZ cities compared to non-TCZ cities.

The dissertation of Mengshan Cui is approved.

Moshe Buchinsky

Adriana Lleras-Muney

George M De Shazo

Dora Luisa Costa, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

VITA

2013	B.A. (Economics and Mathmatics), Brandeis University.
2015	M.A. (Economics), UCLA, Los Angeles, California.
2014–2019	Teaching Assistant, Economics Department, UCLA.

CHAPTER 1

Demand and Supply for Rooftop Solar: Evidence from the California Solar Initiatives

1.1 Introduction

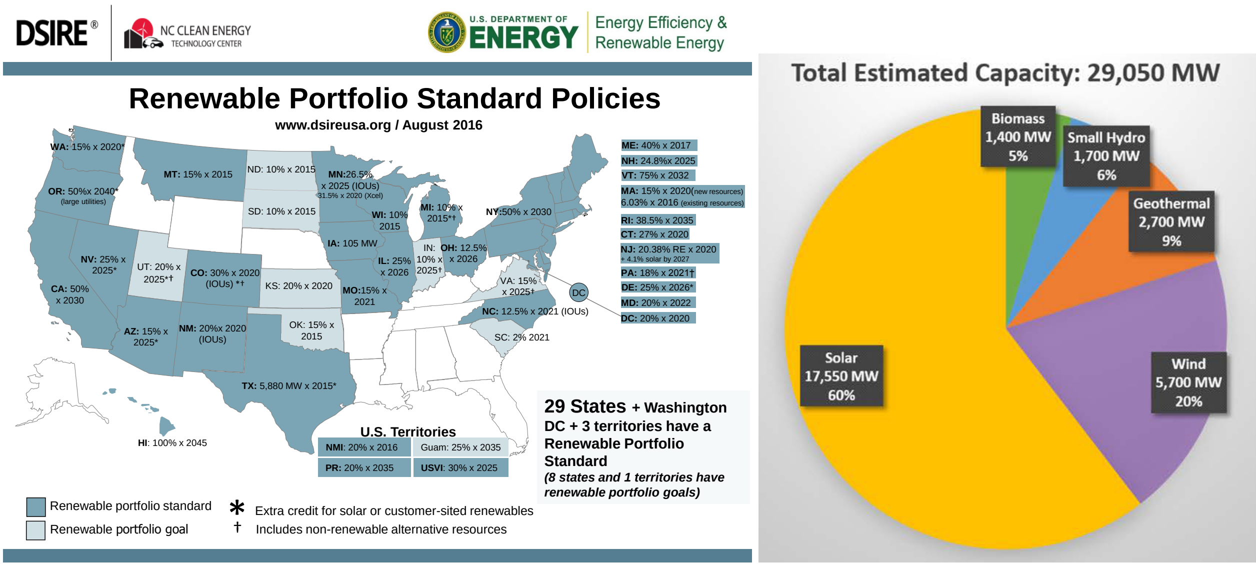
Clean energy has become a hot topic in recent years due to its positive impact on the environment and the economy. In September 2018, California Governor Jerry Brown signed SB100, a bill that requires utility companies to ensure that 60% of the electricity they generate and sell comes from renewable sources by 2030, and 100% by 2045[1].In 2017, 32% of retail electricity came from renewable sources (California Energy Commission Report, 2018). Of the 29,050 megawatts (MW) of renewable capacity at the end of 2017, 60% (17,550 MW) now comes from solar (Figure 1.1 right).

Distributed solar (rooftop solar) capacity has contributed a substantial part, and its had exponential growth in the past ten years (Figure 1.2), reaching 5,793 MW in 2017. A solar rebate program launched in 2007, called the California Solar Initiative (CSI), could have contributed to this exponential increase in distributed solar capacity. The CSI had a budget of \$2.167 billion between 2007 and 2016 and a target to install 1,940 MW of new solar-generation capacity.[2] The CSI's capacity-oriented, one-time rebate program makes it

[1]More than half of U.S. states have established renewable portfolio standards (RPS), which require utility companies to ensure that a percentage of the electricity that they generate and sell comes from renewable resources. By 2016, only 13 states lacked an RPS (Figure 1.1 left). The states with the most stringent RPS are Vermont (75% by 2032), the District of Columbia (50% by 2032), New York (50% by 2030), and Oregon (50% by 2040). California started to establish its RPS in 2002 and accelerated it in 2006 to 20% by 2010.

[2]This program is different from other programs in the United States and around the world in that it offers a one-time cash-back rebate for households that install solar panels on their property. The federal government uses tax credits and European countries use feed-in tariffs, in which governments pay for power generated from household solar panels.

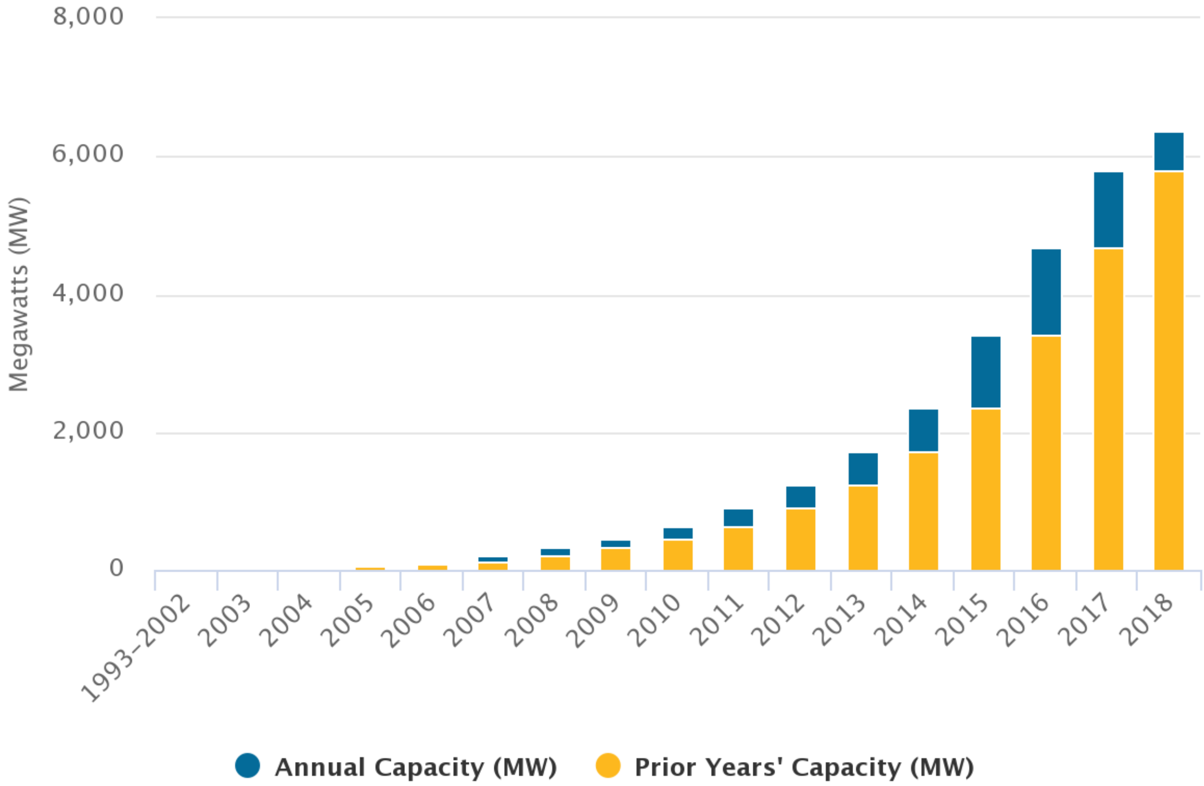
Figure 1.1: Renewable Portfolio Standards and Renewable Energy Capacity in California



possible to analyze the relationship between price and quantity of the solar-panel system and to further study the demand and supply functions of the solar panels. This will help policy makers better understand the effects of the program, such as how many households installed solar panels, how much capacity increased because of the rebate program, and who benefited. Demand and supply functions can also be used to evaluate consumer and producer welfare increases caused by the rebates. Moreover, knowing the market makes it possible to know what kind of programs would be most efficient in achieving future renewable energy targets.

This paper estimates the demand and supply elasticity of rooftop solar-panel systems in California. It is hard to conduct elasticity estimation simply by using observed price and quantity because of simultaneous causality: not only do price changes affect quantity, quantity changes also affect price. To solve the endogenous problem, I use exogeneous changes in rebate and their impact on price and quantity to derive the demand and supply functions. Another difficulty I face in this research: when I analyze disaggregated datasets, for example (at the Zipcode-month level) a large number of Zipcode-months has zero solar-panel adoptions. Disaggregated data is desirable, because it allows me to observe local variations and better control for shocks that shift demand and supply. To solve the excess zero problem, I use a two-part model to incorporate the distribution of zeros. To perform

Figure 1.2: Distributed Solar Capacity in California



the analysis, I combine datasets from multiple sources to extract the information needed. My main dataset is the California Solar Initiative (CSI) Working dataset, which comes from California Distributed Generation Statistics. This application-level dataset contains every application for a rebate under the CSI. My other datasets are the American Community Survey 1% dataset, which has annual county-level demographic variables; solar radiation data from the National Solar Radiation Database; and voting data from the California Statewide Database.

This paper is closely related to literature that studies the impact of rooftop-solar incentive programs in California (Burr (2014); Dong et al. (2018); Rogers and Sexton (2014); Hughes1 and Podolefsky (2015)) and around the rest of the country (Sarzynski et al. (2012); Gillingham and Tsvetanov (2016); Cragoa and Chernyakhovskiy (2017)). Rogers and Sexton (2014) and Hughes and Podolefsky (2015) estimate the impact of rebates on the quantity of rooftop solar-panel systems, while Dong, Wiser, and Rai (2014) study the pass-through rate of rebates using data from the CSI. Meanwhile, Crago and Chernyakhovskiy (2017) use county-level panel data in the Northwest to examine the effectiveness of different financial incentives for increasing solar-panel capacity. Sarzynsk, Larrieu, and Shrimali (2012) use cross-sectional data to study the impact of varies state-level financial incentives on the development of solar-generating capacity. Yang et al. (2003) develop a Bayesian method to address the difficulty of estimating simultaneous demand and supply models. Gillingham and Tsvetanov (2016) is the only paper I find that tries to estimate the demand function while accounting for markets with zero adoptions. They use panel data on the count of annual solar photovoltaic systems installed in a census-block group in Connecticut, using a hurdle model to incorporate zeros. They estimate a price elasticity of demand for solar photovoltaic systems of -1.76. Most of this literature has focused on the effect of rebates on quantity or price. My work goes one step further to derive the supply and demand functions, which will help policy makers better understand the solar-panel market in California and thus design policies accordingly.

I contribute to this literature in two ways. First, my research embodies the first attempt to study both the demand and the supply functions of the California residential solar market. I develop a new way of using rebate shifts and their impact on price and quantity to derive the demand and supply functions. This allows me to estimate demand and supply concurrently—most of the previous literature focused only on either supply or demand, or studied the impact of rebates on quantity or price rather than the relationship between these two. Second, I analyze disaggregated market data at the Zipcode-month level. This allows me to observe and use local variations to control demand and supply shifts. However this creates a problem: that a large part of market would have zero applications for solar-panel systems. Previous literature either defines the market in large aggregates (at the county or annual level) or focuses only on observed positive quantities. However, the number of markets with zero applications becomes nontrivial at the disaggregate level.

I estimate elasticities using two different models. I use a two-part model on data including Zipcode-months with both zero solar-panel applications and positive solar-panel applications.

I also present results using a linear model on data with only Zipcode-months with at least one solar-panel application. I show that these two models lead to different results. A two-part model estimation is preferable because the impact of price on distribution of markets with zero solar-panel applications is important. Using a two-part model, I estimate demand elasticity to be –1.468 and supply elasticity to be 1.451. These estimates for elasticity are higher in absolute value when I focus only on a small period before and after each rebate rate change. When I'm looking only at three months before and after rebate rate changes, demand elasticity is –3.284, while supply elasticity is 5.572. These estimates are more accurate, because focusing on a small period helps to control for unobserved shifts in demand and supply functions. Excluding markets with zero solar-panel applications leads to different elasticity estimations. Using only Zipcode-months with positive applications, I estimate demand elasticity to be –1.507 and supply elasticity to be 0.649. Demand elasticity is similar between the two models, but supply elasticity is more elastic, because when price changes, markets change from zero to nonzero quantity, or vice versa. This cannot be captured by the linear model using only Zipcode-months with at least one application. Using a linear model on data within three months of rebate rate changes, I estimate demand elasticity to be –5.157 and supply elasticity to be 3.882. This is different from the two-part model estimations. Including markets with zero applications can affect elasticity estimation in either direction. Rising prices can either change markets from nonzero applications to zero applications, which would lead the two-part model to have a more elastic elasticity estimation than a linear model, or keep markets with zero applications unchanged, which would lead to a more inelastic elasticity estimation when using the two-part model. Either way, including markets with zero applications is crucial for an accurate estimation of elasticities.

I use demand and supply functions to estimate welfare changes caused by the CSI. Using the estimation of the two-part model, I estimate that the CSI led to a consumer surplus increase of $269 million and a producer surplus increase of $256 million. However, the CSI also caused $663 million of deadweight loss. (This estimation does not account for the positive effects of solar panels on environmental conditions or global warming.) Among the surplus generated, more than 40% went to the top 25% wealthiest Zipcodes, while less than

7% went to the 25% poorest Zipcodes. This implies that the CSI benefits high-income groups most.

I further estimate the impact of rebates on the quantity and price of solar panels. If I use the estimation from my two-part model including Zipcode-months with zero solar-panel applications, I estimate that a \$1 increase in rebate would increase the price received by sellers by \$0.29 and would decrease the price paid by households by \$0.71. Using the estimation from my linear model excluding Zipcode-months with zero solar-panel applications yields different values. Each dollar of rebate increases the price received by sellers by \$0.61 and decreases the price paid by households by \$0.39. This illustrates that using my two-part model to incorporate zero quantity is important in terms of deriving correct evaluations for programs such as the CSI. I also estimate the CSI's impact on consumer behavior and increased power-generating capacity using estimated demand and supply functions. Each dollar increase in the rebate level would increase the number of households that install a solar-panel system by 1.41 per Zipcode each month and increase generating capacity by 7.5 watts per Zipcode each month. Based on each market's rebate level, I estimate that 119,456 (46%) new solar-panel systems and 635 MW (29%) of new capacity were caused by the CSI between 2007 and 2014.

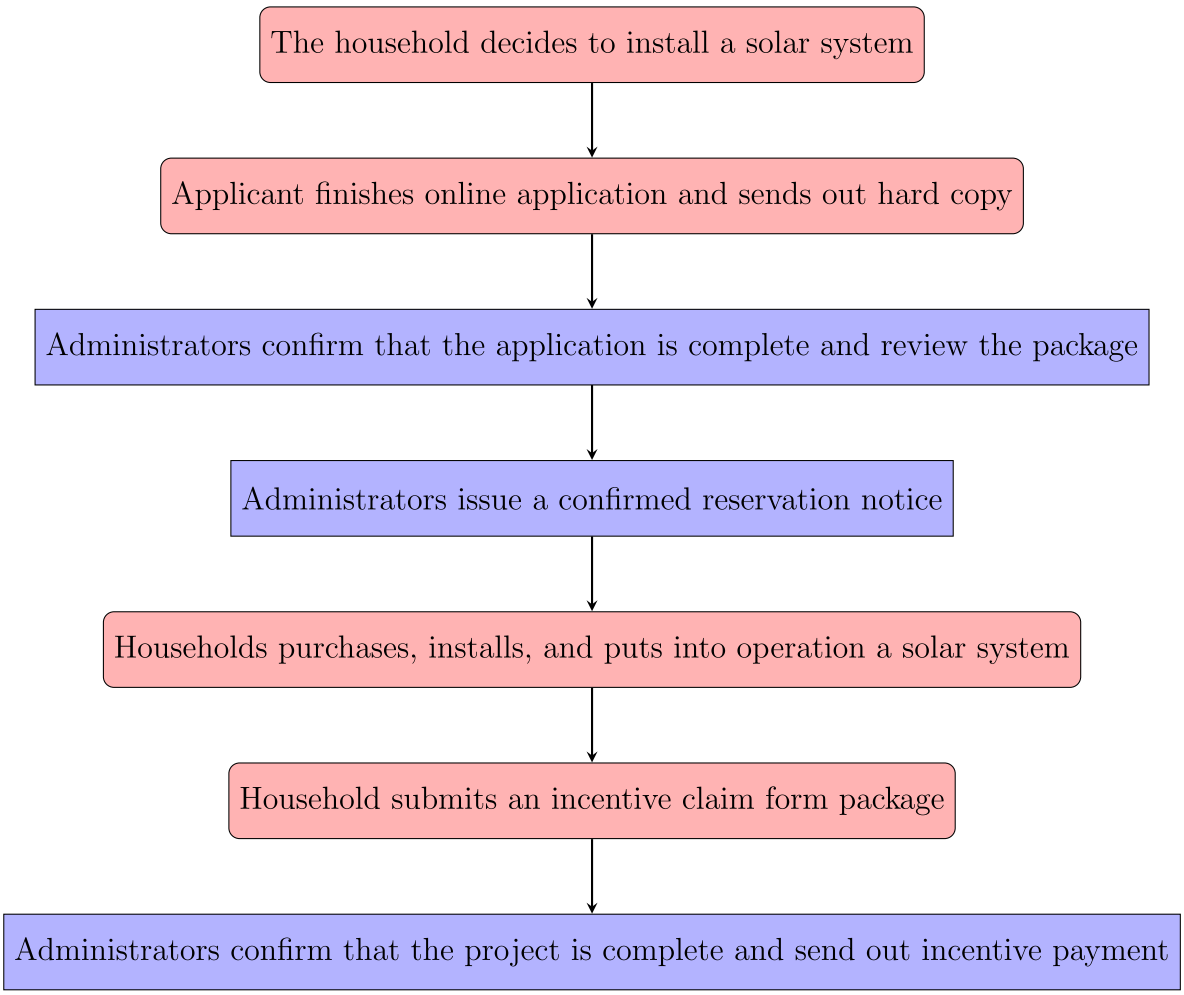
The rest of this paper is organized as follows: Section 1.2 introduces the California Solar Initiatives program. Section 1.3 explains the supply-and-demand model, incorporating rebate rate changes, and further describes my identification methods for analyzing the model. Section 1.4 describes my datasets and provides summary statistics. Section 1.5 presents the specifications and the main results. Section 1.6 provides analysis of the heterogeneous effect. Section 1.7 uses estimated elasticities to calculate the extent of any welfare increase caused by the CSI. Section 1.8 provides a robustness check. Section 1.9 concludes.

1.2 California Solar Initiative

California Solar Initiative (CSI) is the largest solar rebate program in the United States, with a total budget of \$2.167 billion between 2007 and 2016 and a goal to install approximately

1,940 Mw of new solar-generation capacity. It covers the combined utility area served by Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). It funds solar projects on residential, commercial, agricultural, government, and nonprofit buildings. Policies governing each sector are different; in this study, I focus only on the residential sector.

Figure 1.3: Residential Rebate Application Process



Households apply for a solar panel rebate under the CSI (see Figure 1.3) as follows: once a household decides to install new solar panels on their rooftop, they file an online application and mail in proof of their intent to have a solar panel system. This proof can be a contract with a solar contractor or a receipt for a solar system. Once the CSI administrators receive the application and proof, they will process the application. If everything is valid, they will reserve a certain amount of funding within the current step(will explain later) for the household, which depends on the application date and the size of the system. The CSI administrators will send a confirmation notice to the household informing them of the current rebate rate, the confirmed size of the solar-panel system, and total amount of reserved funding for this project. The household then has around 12 months to install the system and connect it to the grid. Once this is done, they can file an incentive claim. After the CSI administrators confirm the system is complete and running, the household will receive a one-time check with the reserved funding.

Interestingly, the CSI rebate rate is not constant; it decreases over time. The CSI first divides the total target capacity into nine steps. Each step is further divided into residential and nonresidential sectors, and finally, the capacity is parceled out to the three utilities. The nine steps happen in sequence. Each utility starts from step 2[3] and when the capacity from step 2 is fulfilled by applications, it will move to step 3. When the step 3 target capacity is reached, it will move to step 4, and so on. The rebate rate decreases with each step. Table 1.1 provides details about each utility's capacity in each step and the rebate rate for each step. For example, in step 2, 70 MW of target capacity is divided into residential and nonresidential sectors. In the residential sector, 10.1 MW is assigned to PG&E, 10.6 is assigned to SCE, and 2.4 is assigned to SDG&E. The rebate rate in step 2 is \$2.50 per watt. This means if a household is applying for the rebate for a 5-kilowatt solar system when the program is in step 2, it will receive a \$12,500(2.5×5000) rebate. The rebate rate drops to \$2.20 per watt in step 3, \$1.90 per watt in step 4, and so on. Furthermore, since each utility gets a different assigned capacity, each utility can move from to step at different times.

This kind of trigger mechanism makes the rebate-rate change exogenous, which means when a household starts to apply for the rebate, they know what the current step is but they do not have control over what step they are going to be in. This will help us both to identify how price and quantity change in response to rebate-rate changes and to determine the demand and supply of rooftop solar panels in California.

[3]The program began with Step 2. Step 1 is reserved for special purposes.

Table 1.1: Incentive Trigger Mechanism

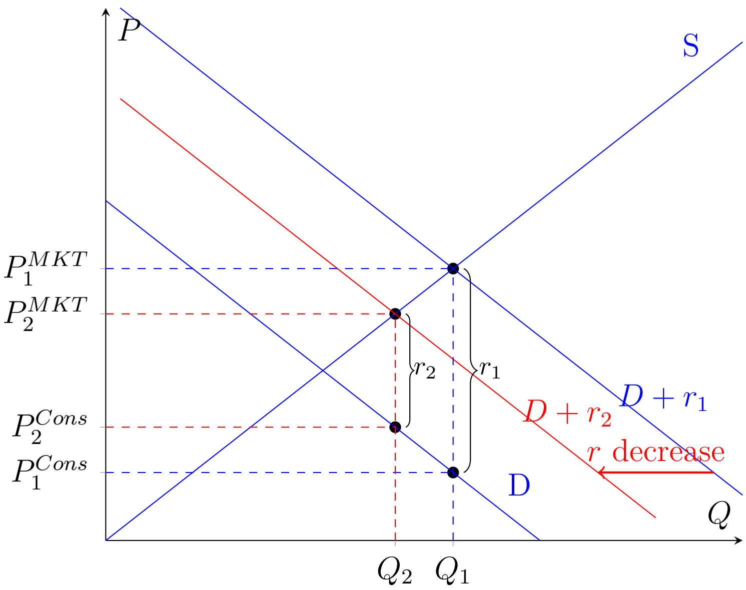
		Residential Sector			Rate
Step	MW in Step	PG&E (MW)	SCE (MW)	SDG&E (MW)	($ per Watt)
2	70	10.1	10.6	2.4	2.5
3	100	14.4	15.2	3.4	2.2
4	130	18.7	19.7	4.4	1.9
5	160	23.1	24.3	5.4	1.55
6	190	27.4	28.8	6.5	1.10
7	215	31.0	32.6	7.3	0.65
8	250	36.1	38.0	8.5	0.35
9	285	41.1	43.3	9.7	0.25
10	350	50.5	53.1	11.9	0.20

1.3 Identification and Methodology

How can one identify the demand and supply elasticities for rooftop solar panel systems in California? This question is hard to answer due to the endogeneity of equilibrium price and quantity. To identify demand elasticity, researchers need to observe several points on the demand function, and to identify supply elasticity, researchers need to observe several points on the supply function. However, most of the time researchers observe only equilibrium points. This would be hard to use unless we can control for the shift in demand and supply functions. In this paper, I argue that we can solve this problem by using exogeneous changes in the rebate rate.

Figure 1.4 depicts a basic supply-and-demand model that assumes a competitive market. Without government rebates, there will be one equilibrium point, where given the equilibrium price the quantity demanded, and the quantity supplied are the same. When a rebate (r_1) is offered, the equilibrium will change. The price received by seller(P_1^{MKT}) will be higher than the price paid by households(P_1^{Cons}) by the amount of the rebate, but the quantity demanded and quantity supplied (Q_1) will be the same. Therefore, given a rebate level, I can observe two points in the market, one on the demand function(Q_1, P_1^{Cons}) and the other on the supply function(Q_1, P_1^{MKT}). When the rebate level changes (to r_2), a new equilibrium will result with one point on the demand function(Q_2, P_2^{Cons}) and another point

Figure 1.4: Basic Supply and Demand Curve



on the supply function(Q_2, P_2^{MKT}). Therefore, if I look only at a short period before and after the rebate change, I can assume that the demand and supply functions are fixed and I can observe multiple points on both functions, which will allow me to identify the demand and supply elasticities.

Above I describe my intuition and justification for the identification methodology I use in this paper. I provide the mathematical model later in this section. Unfortunately, this simple model considers the supply and demand only given there are positive number of household buying solar panel systems. In real life, because rooftop solar panel is expensive and not necessary for every single household, in a lot of cases there is no transaction in the market, especially when the market is finite (for example, Zipcode-month). I provide more summary statistics about excess zero in the summary statistics section. But I also need statistical model to incorporate the excess zeros.

1.3.1 Two-part Model by Duan et al. (1983)

The two-part model presented in Cragg (1971) and Duan et al. (1983) can be used to solve excess zero and model continuous data. I define a binary indicator variable $d = 1$ for observed quantity and $d = 0$ for zero quantity. For zeros, I observe $Pr(d = 1)$, while for observed quantity the conditional density for q given $q > 0$ is specified as $f(q|d = 1)$. Therefore, the two-part model for quantity is given as follows:

$$f(q|x) = \begin{cases} Pr[d = 0|x] & \text{if } q = 0 \\ Pr[d = 1|x]f(q|d = 1, x) & \text{if } q > 0. \end{cases} \tag{1.1}$$

A latent variable formulation is used to model the zero generating process: $d = 0$ if $I = x_1'\beta_1 + u_1$ exceeds zero. This is called a hurdle model, since crossing a threshold leads to zero. I specify the following probit model for this process:

$$d = \begin{cases} 0 & \text{if} \quad x_1'\beta_1 + u_1 > 0 \\ 1 & \text{if} \quad x_1'\beta_1 + u_1 < 0, \end{cases} \tag{1.2}$$

where u_1 follows the standard normal. If all x_1 are exogenous,

$$Pr[d = 0|x] = \Phi(x_1'\beta_1). \tag{1.3}$$

Conditional on a quantity not equal to zero, I use a log-normal distribution to ensure that the quantity is positive:

$$log(q|d = 1) = x_2'\beta_2 + u_2, \tag{1.4}$$

where $u_2 \sim N(0, \sigma^2)$. Hence, the likelihood function is

$$\begin{aligned} l(\beta_1, \beta_2, \sigma) &= \Big[\prod_{q_i=0} P(d_i = 0)\Big]\Big[\prod_{q_i>0} P(d_i = 1)f(q_i|d_i = 1, x_i)\Big] \\ &= \Big[\prod_{q_i=0} \Phi(x_1'\beta_1)\Big]\Big[\prod_{q_i>0} \Big(1 - \Phi(x_1'\beta_1)\Big)\sigma^{-1}\phi\Big(\frac{log(q_i) - x_{2i}'\beta_2}{\sigma}\Big)\Big]. \end{aligned}$$

This can be factored into $l_1(\beta_1)$ and $l_2(\beta_2, \sigma)$, as follows

$$l_1(\beta_1) = \Big[\prod_{q_i=0} \Phi(x_1'\beta_1)\Big]\Big[\prod_{q_i>0}\Big(1 - \Phi(x_1'\beta_1)\Big)\Big] \tag{1.5}$$

and

$$l_2(\beta_2, \sigma) = \Big[\prod_{q_i>0} \sigma^{-1}\phi\Big(\frac{log(q_i) - x_{2i}'\beta_2}{\sigma}\Big)\Big]. \tag{1.6}$$

Therefore, I can obtain maximum likelihood estimators by separately maximizing $l_1(\beta_1)$ and $l_2(\beta_2, \sigma)$.

Endogenous Price

I also need to consider the endogeneity of price. I exploit the exogeneity of rebates. Two-stage least square estimators do not work well in nonlinear models. Control function estimators perform better in this setup. With endogeneity of price, the model is changed:

$$d = \begin{cases} 0 & \text{if} \quad \theta_1 Price + x_1'\beta_1 + u_1 > 0 \\ 1 & \text{if} \quad \theta_1 Price + x_1'\beta_1 + u_1 < 0, \end{cases} \tag{1.7}$$

where $\theta = \delta$ for demand and $\theta = \gamma$ for supply. Since price is endogenous,

$$Price = \pi Rebate + x'\beta + v, \tag{1.8}$$

where I assume

$$v | Rebate, x \sim N(0, \sigma_v^2) \tag{1.9}$$

$$u_1, v \perp\!\!\!\perp Rebate, x \tag{1.10}$$

$$(u_1, v) \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_v\rho_1 \\ \sigma_1\sigma_v\rho_1 & \sigma_v^2 \end{pmatrix}\right]. \tag{1.11}$$

since u_1 and v are joint normal, the conditional probability of $u_1|v$ is

$$u_1|v \sim N\Big(\frac{\rho_1}{\sigma_v}v, (1-\rho_1^2)\Big), \tag{1.12}$$

and

$$\begin{aligned}
Pr(d=0) &= Pr(\theta_1 Price + x_1'\beta_1 + u_1 > 0) && (1.13)\\
&= Pr(u_1 > -(\theta_1 Price + x_1'\beta_1)) && (1.14)\\
&= Pr\left(\frac{u_1 - \frac{\rho_1}{\sigma_v}v}{\sqrt{1-\rho_1^2}} > -\frac{\theta_1 Price + x_1'\beta_1 + \frac{\rho_1}{\sigma_v}v}{\sqrt{1-\rho_1^2}}\right) && (1.15)\\
&= \Phi\left(\frac{\theta_1}{\sqrt{1-\rho_1^2}}Price + x_1'\frac{\beta_1}{\sqrt{1-\rho_1^2}} + \frac{\rho_1/\sigma_v}{\sqrt{1-\rho_1^2}}v\right). && (1.16)
\end{aligned}$$

When using a probit model to fit price, controls, and estimated v, the coefficients need to be adjusted to back out the original ones. I follow these steps:

Step 1. Estimate

$$Price = \pi Rebate + x'\beta + v$$

and get $\hat{v}$ and $\hat{\sigma_v}$.

Step 2. Fit the probit model on price, control, and $\hat{v}$.

Step 3. Adjust the coefficient by using

$$1 + \beta_v^2\hat{\sigma_v}^2 = \frac{1}{1-\rho_1^2}$$

, where β_v is the preadjusted estimated coefficient on $\hat{v}$. Hence,

$$\theta_1 = \beta_P \times (1 + \beta_v^2\hat{\sigma_v}^2)^{-1/2}$$

$$\beta_1 = \beta_x \times (1 + \beta_v^2\hat{\sigma_v}^2)^{-1/2}$$

, where β_P and β_x are preadjusted estimated coefficients on $Price$ and X_1.

The positive log-normal model also needs to be adjusted to identify the impact of price on positive quantity:

$$log(q) = \theta_2 Price + x_2\beta_2 + u_2. \tag{1.17}$$

Because of endogeneity, I assume

$$u_2 = \rho_2 v + w, \tag{1.18}$$

where $w \sim N(0, \sigma_w^2)$. Plug the followings into equation 1.17:

$$log(q) = \theta_2 Price + x_2\beta_2 + \rho_2 v + w. \tag{1.19}$$

To identify θ_2 in this model, I first use Step 1 above to get $\hat{\sigma_v}$, then I use it as a control to fit the above equation.

Expected Quantity

Under the two-part model, the expected size of a new solar system for a Zipcodes in a month would be

$$E(q_i|x_i) = (1 - \Phi(x_{1i}\beta_1)) \times exp\{\frac{\sigma^2}{2}\} \times exp\{x'_{2i}\beta_2\}$$

. I will use this equation later to estimate the impact of price on quantity.

1.3.2 Using Only Zipcode-month with at Least One Solar Panel Application

If I consider only Zipcode-month with at least one solar panel application, I can use a linear demand-and-supply model and equilibrium conditions to model the solar panel markets.

Demand:

$$Q^d = \alpha^d + \delta P^{Cons} + \epsilon^d \tag{1.20}$$

Supply:

$$Q^s = \alpha^s + \gamma P^{MKT} + \epsilon^s, \tag{1.21}$$

where Q^d and Q^s are demand quantity and supply quantity, P^{Cons} and P^{MKT} are price paid by consumers and price received by sellers. I am interested in identifying δ, which is demand elasticity, and γ, which is supply elasticity. Therefore, the equilibrium conditions are:

$$Q^d = Q^s \qquad \Longrightarrow \qquad Q^d = Q^s \tag{1.22}$$

$$P^{Cons} + rebate = P^{MKT} \tag{1.23}$$

$$\tag{1.24}$$

I combine supply, demand, and equilibrium conditions:

$$Q - \delta P^{MKT} = \alpha^d - \delta rebate + \epsilon^d \tag{1.25}$$

$$Q - \gamma P^{MKT} = \alpha^s + \epsilon^s \tag{1.26}$$

Next, I rewrite the above equations in matrix:

$$\begin{bmatrix} 1 & -\delta \\ 1 & -\gamma \end{bmatrix} \begin{bmatrix} Q \\ P^{MKT} \end{bmatrix} = \begin{bmatrix} \alpha^d & -\delta \\ \alpha^s & 0 \end{bmatrix} \begin{bmatrix} 1 \\ rebate \end{bmatrix} + \begin{bmatrix} \epsilon^d \\ \epsilon^s \end{bmatrix}.$$

Finally, I solve

$$\begin{aligned} Q &= \alpha^q + \frac{\delta\gamma}{\delta - \gamma} rebate + \epsilon^q \\ P^{MKT} &= \alpha^p + \frac{\delta}{\delta - \gamma} rebate + \epsilon^p. \end{aligned} \tag{1.27}$$

Equation 1.27 implies that if I can identify the impact of a rebate on the quantity and market price of solar panels, I can identify demand elasticity δ and supply elasticity γ.

1.4 Dataset and Summary Statistics

1.4.1 Data Sources

My main dataset is the California Solar Initiative (CSI) Working dataset, which comes from California Distributed Generation Statistics. This application-level dataset contains all applications for rebates under CSI, including time of the application, the location of the property (Zipcodes), the CSI step (as described in the Introduction), the size of the solar panel, total cost of the solar system, and the total rebate received. To perform market-level analysis, I aggregated application-level data into the Zipcode-month level. The month is based on the time when the application is accepted and being reviewed, which I consider as the time when a household decided to install solar panels, and the CSI step is determined by the time of the application.

My other datasets include the American Community Survey 1% dataset, which contains annual county-level demographic variables such as household income, education, and unemployment. Solar-radiation data from the National Solar Radiation Database has station-level hourly data. I use this to calculate solar-radiation intensity at the Zipcode-month level. This is important because households in high-solar-radiation areas can see higher returns their on solar panels. Voting data come from the California Statewide Database. At the voting-precinct level for each voting year. I use the overlap between voting precinct and Zipcodes to estimate number of registered Democrats and Republicans. Political opinion can indicate local preference for environment-friendly products.

For Zipcode-months with zero applications, I do not observe price at the Zipcodes for that month. Therefore, I need to estimate the possible price of solar panels. I use the price level in nearby Zipcodes or from the same county as a proxy for missing prices in the Zipcode-months with zero applications. Rebate levels for these Zipcodes are based on the CSI step.

1.4.2 Summary Statistics

Table 1.2 provides summary statistics for my main Zipcode-level variables using all Zipcode-months and using only Zipcode-months with at least one applications. Among 76,576 Zipcode-month observations, only 35,393 of them have positive amount of solar panel system application. Given that some households apply for the rebate, the number of applications per Zipcode-month is on average around three. The total capacity of applications per Zipcode-month is more than 7 kilowatts(kW) if I use all Zipcode-months and around 16.5kW if I use

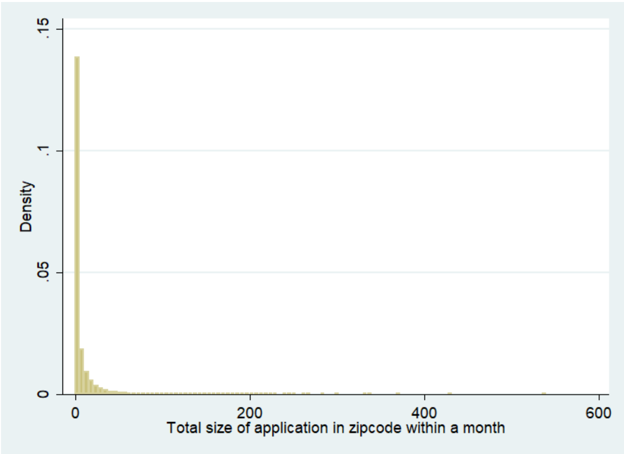
Table 1.2: Summary Statistics of Important Variables

	All Zipcode-Month		Only Zipcode-Month with Nonzero Applications	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Obervations	76,576		35,393	
Total number of applications	1.303006	2.58241	2.819173	3.186598
Total size(Kw)	7.634052	16.85783	16.51697	21.63677
Market price($/watt)	6.925683	1.684893	6.629399	1.728156
Price paid by consumers($/watt)	5.820991	1.274974	5.82254	1.364886
Rebate ($/watt)	1.104862	.8390347	.8068597	.6721016
House price variables:				
Average house price($1000)	475.45	347.62	495.43	363.52
Average single family house price($1000)	509.21	392.95	529.99	404.35
Average price per square feet	280.8998	156.5137	280.3601	159.027
Voting registration variables:				
Total number of registrations	14089.27	8537.318	16205.51	8182.841
Number of Democrats	5855.664	3955.332	6460.943	3766.753
Number of Republicans	4757.69	3658.732	5702.375	3676.768
Ratio of Republicans	.3448191	.1340634	.3561128	.1281744
Ratio of Democrats	.411692	.1185487	.3974852	.1119328

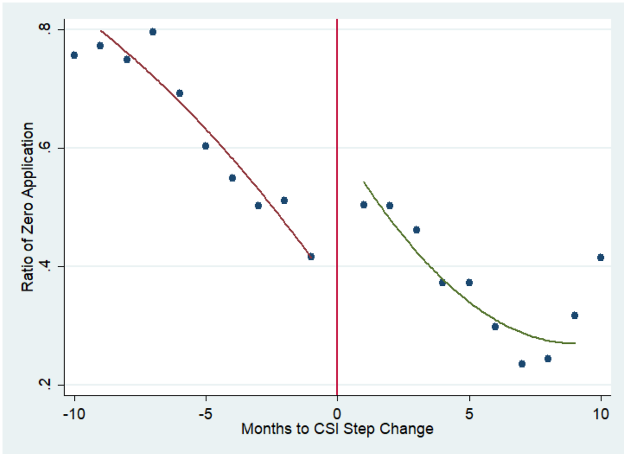
only Zipcode-month with at least one application. Figure 1.5(a) provides a histogram distribution of the Zipcode-month-level capacity of applications. There is a high concentration at zero. The distribution is high at the low-level capacity and has a long tail. Figure 1.6 compares the number of Zipcodes with zero applications with the total applications for each month. The number of months with zero applications fluctuates, but it overall decreases from 2007 to 2014. The trend for number of total applications has the opposite pattern. This indicates that the factor that affects the number of zero applications also affects the number of total applications, with different signs. Figure 1.5(b) presents the ratio of Zipcodes with zero quantity over all Zipcodes around the time of the CSI-step change. Clearly, the trend shifts up after the rebate decreases. This is evidence that when the rebate rises, the price paid by households increases, which affects the probability of zero quantity. Hence, I must incorporate zero quantity in my market analysis.

The price for a rooftop solar panel system is similar when using all Zipcode-months and using only Zipcode-months with at least one application. The market price is a little less

Figure 1.5: Zero Distribution

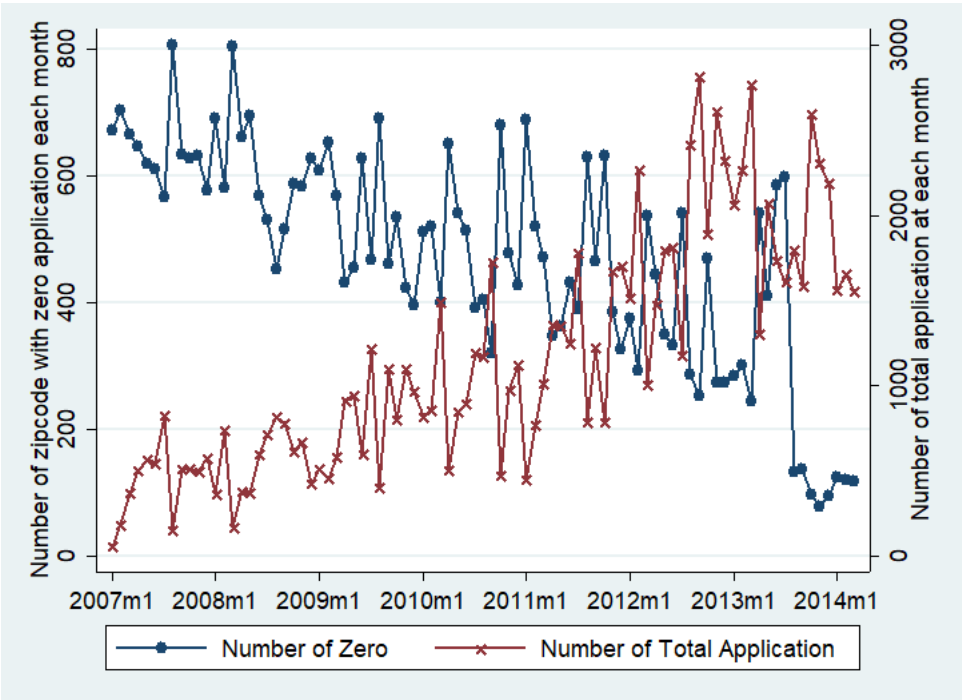


(a) Zipcode-month Application Capacity Distribution



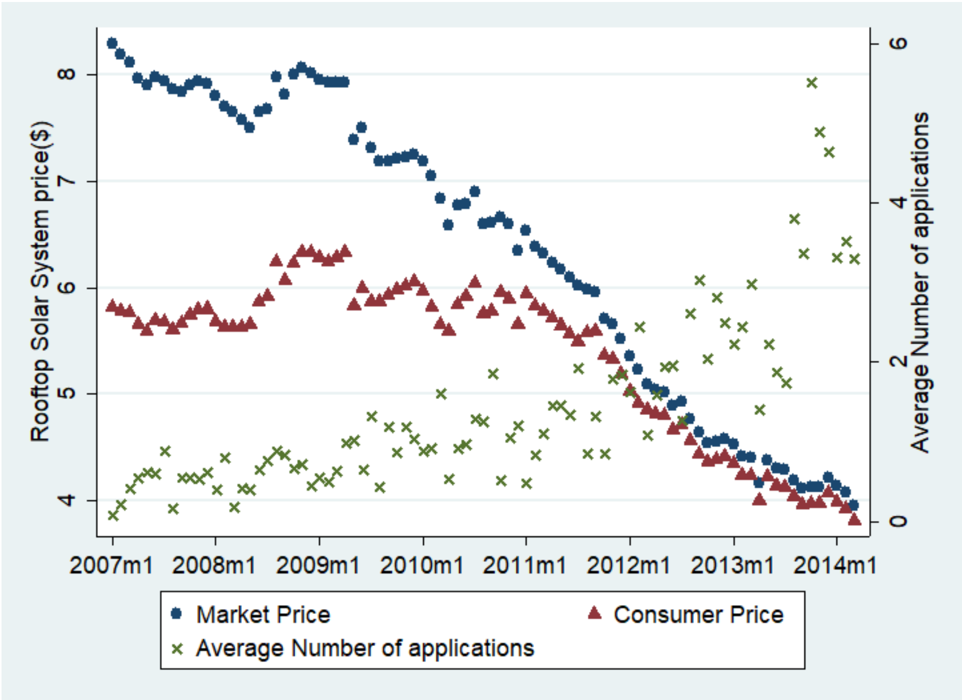
(b) Ratio of number of Zipcodes with zero quantity to total number of Zipcodes around CSI-step change

Figure 1.6: Number of Zipcodes with zero applications compared with total number of applications



than \$7 per watt and the price paid by households is a little less than \$6 per watt. Figure 1.7 displays the price trend of solar panels from January 2007 to March 2014. (Data after March

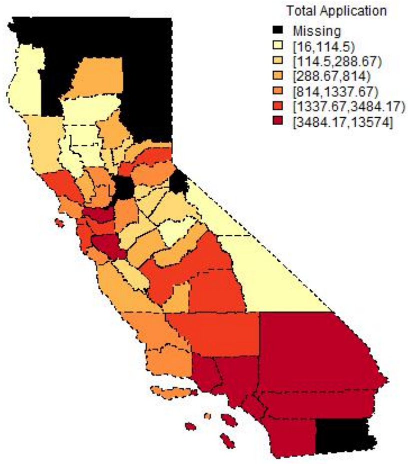
Figure 1.7: Price and Application Trends



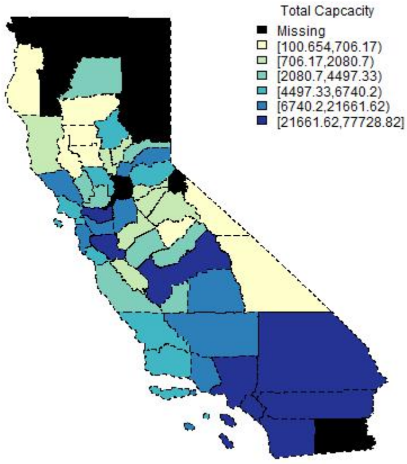
2014 is not included since the program capacity ran out around that time.) The average market price for rooftop solar panels continuously decreases from more than \$8 per watt in early 2007 to less than \$4 per watt in March 2014. The gap between market price and consumer price is the rebate paid by the CSI. From 2007 to 2011, the price paid by households is flat, around \$6 per watt, then it starts to decrease quickly. The number of applications continuously increases over the whole period. In this figure, I calculate the average number of Zipcode-level applications using both Zipcodes with at least one application and Zipcodes with zero applications. Since every Zipcodes's step-switch time is different, I cannot provide this time on this figure.

Figures 1.8(a) and 1.8(b) present the county map of California with the total number of applications and the total capacity. These two figures have similar information. Southern California has the highest concentration of households with solar panels and also the highest capacity. Middle West California has slightly lower quantity. Northern California and Eastern California have the fewest solar panels. This distribution is consistent with population distribution. The Bay Area, Los Angeles, and Orange County have most; Northern Cali-

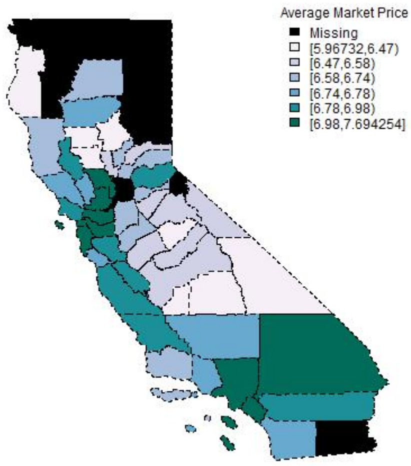
Figure 1.8: County Quantity and Price Distribution



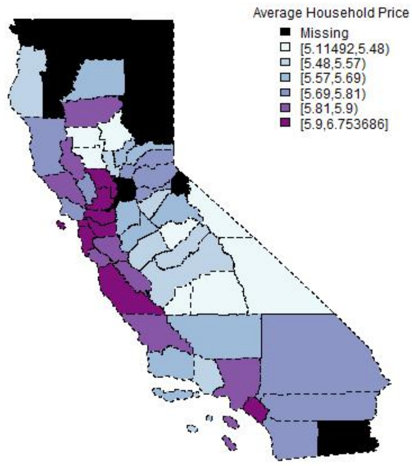
(a) County-Level Distribution of Total Number of Applications



(b) County-Level Distribution of Total Capacity of Applications



(c) County-Level Distribution of Average Price($/watt) of Rooftop Solar-Panel System received by sellers



(d) County-Level Distribution of Average Price($/watt) of Rooftop Solar-Panel System paid by Households fornia and Eastern California have the least. Figures 1.8(c) and 1.8(d) show the geographic distribution of market price and price paid by households. The market price and the price paid by households have similar (but slightly different) distribution. Southern California and the Bay Area have the highest prices. Other areas have lower prices.

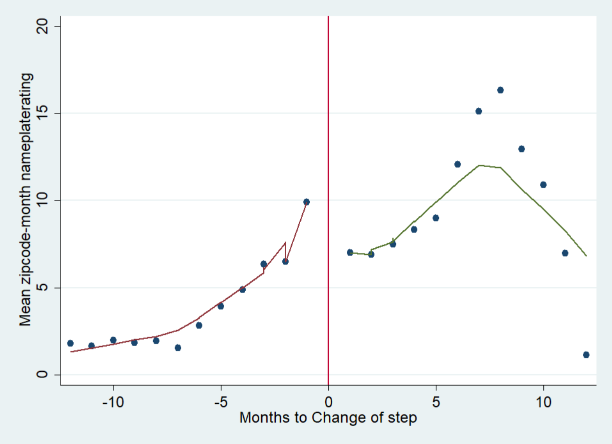
Figure 1.9 show the size of rooftop-solar-system application capacity,the per watt price paid by households, and the per watt price received by sellers before and after a CSI-step change. Figure 1.9(a), Figure 1.9(c), and Figure 1.9(e) show 12 months before and after the step change, while Figure 1.9(b), 1.9(d), and 1.9(f) focus on six months before and after. From these figures, I can observe a clear change in trends for the new capacity of Zipcode-month solar panels and the price paid by consumers. Both are consistent with the theory: when rebates decrease, the quantity will decrease and the price paid by households will increase. However, I do not observe a clear change for the market price. This could be due to price stickiness.

Table 1.3 provides summary statistics for the months between steps. On average, there are eight months between steps. On average, Zipcodes has more than 10 months in Step 2, around 9 months in step 3 and 4, and around seven months for step 5 to 9. There are less households interested in solar panels when the CSI just started in early 2007 because price were higher.

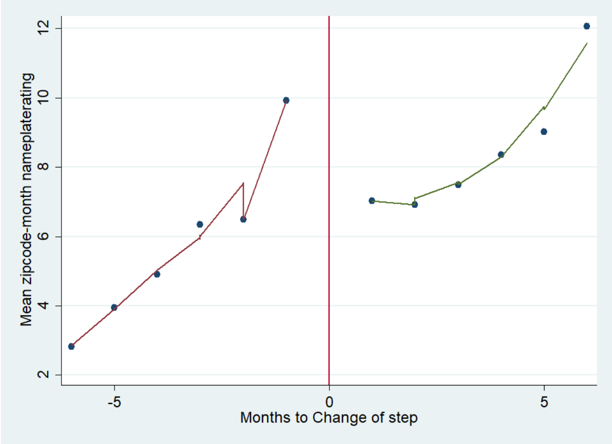
Table 1.3: Gap Between Steps

	Mean	Std. Dev.
Month between all Steps	8.650676	2.707189
Month between steps 2 and 3	10.87912	4.183372
Month between steps 3 and 4	9.42519	3.025991
Month between steps 4 and 5	9.721048	1.803819
Month between steps 5 and 6	7.761623	1.494711
Month between steps 6 and 7	7.401522	.8580648
Month between steps 7 and 8	6.362637	.7156863
Month between steps 8 and 9	9.948436	2.446354
Month between steps 9 and 10	7.705833	1.551602

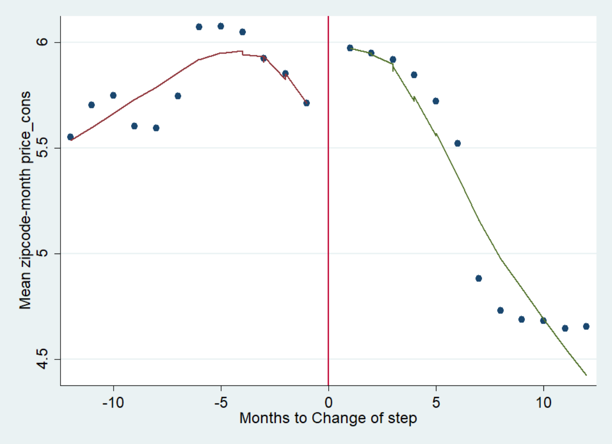
Figure 1.9: Before and After Step Switch



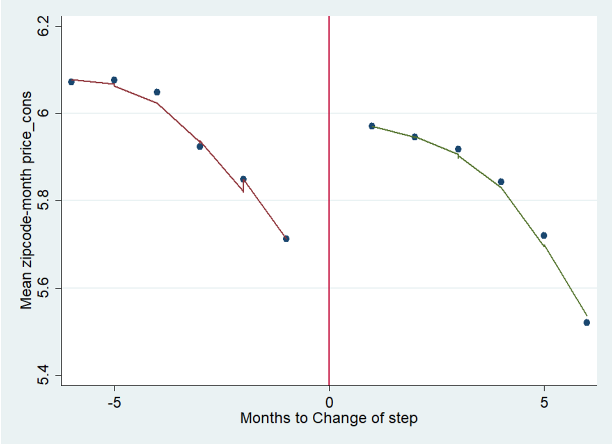
(a) Size of Solar System Before and After Step Switch



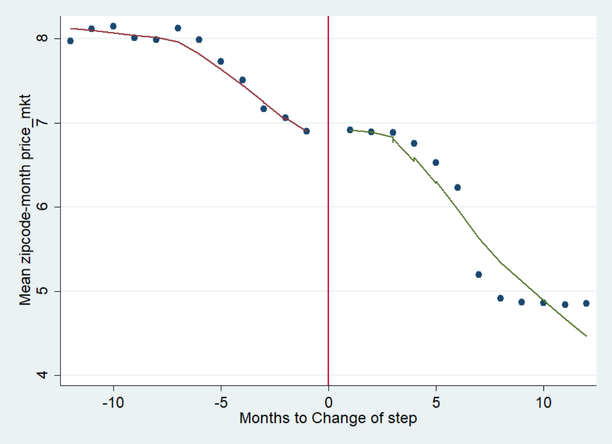
(b) Size of Solar System Six Months Before and After Step Switch



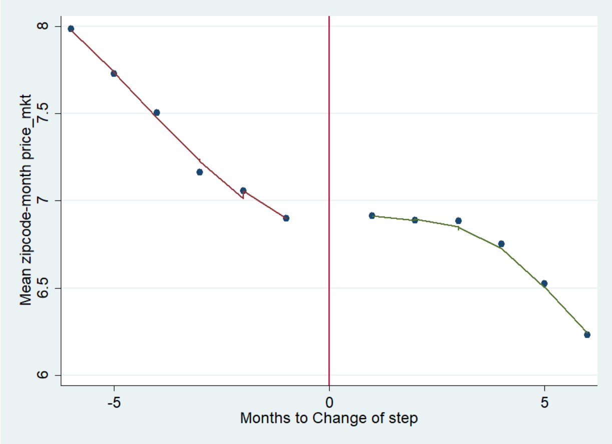
(c) Price Paid by Households Before and After Step Switch



(d) Price Paid by Households Six Months Before and After Step Switch



(e) Price Received by Sellers Before and After Step Switch



(f) Price Received by Sellers Six Months Before and After Step Switch

1.5 Specifications and Results

1.5.1 Two-Part Model by Duan et al. (1983)

Since the two-part model is nonlinear, I use the rebate rate change as an exogeneous shock for price and I use a control function estimator. I first regress market price and price paid by households on the rebate, then I include the estimated residual as an additional control. In Section 1.3, I showed that the estimation for the two-part model can be separated into two parts: predicting zero and fitting positive quantity. The specification proceeds according to the following steps:

Step 1. Regress price on the rebate:

$$P_{it}^{cons} = \pi_0^D + \pi_1^D Rebate_{it} + \pi_2^D X_{it} + v_{it}^D \tag{1.28}$$

$$P_{it}^{mkt} = \pi_0^S + \pi_1^S Rebate_{it} + \pi_2^S X_{it} + v_{it}^S. \tag{1.29}$$

In this step, obtain residual $\hat{v_{it}^D}$ and v_{it}^S and the estimated variance for residuals: $\hat{\sigma_v^D}$ and $\hat{\sigma_v^S}$.

Step 2. Fit the probit model using price, control, and $\hat{v}$:

$$\text{Demand:} \quad Pr(d_{it} = 0) = \Phi\Big(\alpha_{1BA}^D + \delta_{1BA} Price_{it}^{cons} + \beta_{1BA}^D X_{1it} + \beta_{v_{1BA}}^D \widehat{v_{it}^D}\Big) \tag{1.30}$$

$$\text{Supply:} \quad Pr(d_{it} = 0) = \Phi\Big(\alpha_{1BA}^S + \gamma_{1BA} Price_{it}^{mkt} + \beta_{1BA}^S X_{1i} + \beta_{v_{1BA}}^S \widehat{v_{it}^S}\Big). \tag{1.31}$$

Step 3. Adjust the coefficients:

$$\beta = \beta_{BA} \times (1 + \widehat{\beta_{v_{1BA}}} \hat{\sigma_v}^2)^{-1/2}, \tag{1.32}$$

where β is all coefficients. β_{BA} is preadjusted coefficients. $\beta_{v_{1BA}}$ is the preadjusted coefficient on $\hat{v_{it}}$.

Step 4. Fit the positive amount:

$$\text{Demand:} \quad log(q_{it}|q_{it} > 0) = \alpha_2^D + \delta_2 Price_{it}^{cons} + \beta_2^D X_{2it} + \theta_2^D \widehat{v_{it}^D} \tag{1.33}$$

$$\text{Supply:} \quad log(q_{it}|q_{it} > 0) = \alpha_2^S + \gamma_2 Price^{mkt} + \beta_2^S X_{2i} + \theta_2^S \widehat{v_{it}^S}, \tag{1.34}$$

where $Price^{cons}$ represents the price of solar panels that households paid out of pocket, and $Price^{mkt}$ represents the price of solar panels received by sellers.X_1 and X_2 can be different, but I set them to be the same set of controls. I include only Zipcodes level controls in my specification. Location and time fixed effects are included.

Table 1.4: Two-Part model (Dependent Variable: $Price$)

	Demand Estimation			Supply Estimation		
	(1)	(2)	(3)	(4)	(5)	(6)
	$Price^{cons}$	Predict Zero	$ln(size)$	$Price^{MKT}$	Predict Zero	$ln(size)$
$Price^{cons}$		6.471***	0.167***			
		(0.064)	(0.048)			
$Price^{MKT}$					−7.759***	−0.200***
					(0.077)	(0.058)
$\hat{V}$		−6.400***	−0.256***		7.829***	0.110*
		(0.064)	(0.048)		(0.077)	(0.058)
Rebate	−0.545***			0.455***		
	(0.021)			(0.021)		
Previous count	0.001***	−0.003***	0.010***	0.001***	0.013***	0.010***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Total voting	0.001	0.066***	0.039***	0.001	0.079***	0.039***
Registration	(0.005)	(0.008)	(0.005)	(0.005)	(0.008)	(0.005)
Average house	0.096***	−0.679***	−0.103***	0.096***	0.686***	−0.068***
price(100,000)	(0.018)	(0.028)	(0.018)	(0.018)	(0.028)	(0.019)
House price per	−0.318***	2.189***	0.281***	−0.318***	−2.336***	0.164***
Sq.f.t(100,000)	(0.035)	(0.060)	(0.041)	(0.035)	(0.060)	(0.042)
Solar intensity	0.002***	−0.010***	−0.000	0.002***	0.016***	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Ratio of Democrats	0.183	−3.557***	−0.362	0.183	−0.957	−0.295
	(0.538)	(0.833)	(0.567)	(0.538)	(0.832)	(0.566)
Constant	3.862***	−25.554***	1.212***	3.862***	29.400***	2.628***
	(0.214)	(0.412)	(0.309)	(0.214)	(0.434)	(0.333)

Standard errors are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Tables 1.4 and 1.5 display the results of the two-part model. Table 1.4 uses price as the control while Table 1.5 uses $ln(price)$ as the control. The impact of the rebate on price is displayed in Columns (1) and (4) in both tables. This results indicate that a \$1 increase

Table 1.5: Two-Part model (Dependent Variable: $ln(Price)$)

	Demand Estimation			Supply Estimation		
	(1)	(2)	(3)	(4)	(5)	(6)
	$ln(Price^{C})$	Predict Zero	$ln(size)$	$ln(Price^{M})$	Predict Zero	$ln(size)$
$ln(Price^{cons})$		33.450***	–1.652***			
		(0.415)	(0.290)			
$ln(Price^{MKT})$					–54.647***	2.488***
					(0.678)	(0.474)
$\hat{V}$		–33.002***	1.232***		54.865***	–2.994***
		(0.416)	(0.291)		(0.680)	(0.474)
Rebate	–0.066***			0.040***		
	(0.003)			(0.002)		
Previous count	0.000***	0.000	0.010***	0.000***	0.018***	0.009***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Total voting	0.002**	0.056***	0.041***	–0.000	0.118***	0.038***
registration	(0.001)	(0.008)	(0.005)	(0.001)	(0.008)	(0.005)
Average house	0.014***	–0.531***	–0.062***	0.012***	0.599***	–0.115***
price(100,000)	(0.003)	(0.026)	(0.018)	(0.002)	(0.026)	(0.019)
House price per	–0.047***	1.711***	0.137***	–0.039***	–1.961***	0.308***
Sq.f.t(100,000)	(0.006)	(0.056)	(0.041)	(0.005)	(0.057)	(0.042)
Solar intensity	0.000***	–0.009***	0.000	0.000***	0.017***	–0.001*
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Ratio of Democrats	0.103	–3.545***	0.025	0.166**	9.014***	–0.583
	(0.086)	(0.786)	(0.571)	(0.073)	(0.794)	(0.573)
Constant	1.219***	–46.267***	4.094***	1.425***	72.390***	–1.470**
	(0.035)	(0.659)	(0.471)	(0.030)	(0.947)	(0.676)

Standard errors are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

in rebate increases the market price by \$0.455 and decreases the price paid by households by \$0.545. A 1% increase in the rebate increases the market price by 0.04% and decreases the price paid by households by 0.66%. Column (2) and (5) in both tables show how the price change affects the distribution of zero quantity. Both tables demonstrate consistent results: an increase in price paid by households increases the number of Zipcodes with zero applications and an increase in market price decreases the number of Zipcodes with zero applications. However, when I focus only on positive quantity, the relationship is not what I expect when using price as the control. This can be due to markets with a small number of dropouts when the rebate decreases. When I use the $ln(Price)$ as the control, the results indicate that a 1% increase in price would decrease the quantity demanded by 1.652% and increase the quantity supplied by 2.488%, given that the quantity is positive.

Table 1.6 contains the results of using a certain month before and after a CSI-step

Table 1.6: Two-Part Model Results by Months around Step Switch

	Months	All	2 Months	3 Months	4 Months	5 Months	6 Months
	Coefficient on *Price*						
Demand	δ_1	0.930***	0.906***	0.905***	0.893***	0.893***	0.895***
		(0.009)	(–0.006)	(–0.005)	(0.004)	(0.005)	(0.005)
	δ_2	0.167***	–0.132**	–0.168***	–0.100**	–0.066	–0.03
		(0.048)	(0.056)	(0.055)	(0.05)	(0.05)	(0.044)
Supply	γ_1	–0.914***	–0.912***	–0.905***	–0.890***	–0.888***	–0.889***
		(0.009)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)
	γ_2	–0.200***	0.861*	0.412***	0.199*	0.121	0.053
		(0.058)	(0.486)	(0.142)	(0.103)	(0.098)	(0.077)
	Coefficient on $ln(Price)$						
Demand	δ_1	5.812***	5.709***	5.705***	5.663***	5.654***	5.672***
		(0.072)	(0.032)	(0.029)	(0.027)	(0.021)	(0.026)
	δ_2	–1.652***	–3.258***	–3.725***	–3.418***	–3.165***	–2.902***
		(0.290)	(0.360)	(0.338)	(0.364)	(0.331)	(0.342)
Supply	γ_1	–6.783***	–5.453***	–6.743***	–6.683***	–6.675***	–6.702***
		(0.084)	(3.989)	(0.031)	(0.030)	(0.026)	(0.029)
	γ_2	2.488***	112.005	18.632***	12.455***	9.020***	8.313***
		(0.474)	920.434	(4.210)	(2.560)	(1.484)	(1.450)

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

switch. Predicting distribution of Zipcode-months with zero applications is consistent for both specifications, using price and $ln(price)$, and for demand and supply. Fitting positive quantity into the model is also consistent across months for both specifications. This table strongly demonstrates that when households pay more, the number of Zipcodes with zero applications increases and average positive capacity decreases. Conversely, an increase in market price for sellers decreases the number of Zipcodes with zero applications and increases average positive capacity.However, it is hard to use this result to inform policy making without further information about elasticity.

1.5.2 Using Only Zipcode-months with at Least One Rooftop Solar Panel Application

Specification:

$$
\begin{aligned}
Q_{lt} &= \alpha^q + \beta^q Rebate_{lt} + \theta^q X_{lt} + \zeta_t^q + \eta_l^q + \sum_{d=1}^{K} \iota_d^q I_{ltd} + \epsilon_{lt}^q \\
P_{lt}^{MKT} &= \alpha^p + \beta^p Rebate_{lt} + \theta^p X_{lt} + \zeta_t^p + \eta_l^p + \sum_{d=1}^{K} \iota_d^p I_{ltd} + \epsilon_{lt}^p,
\end{aligned}
$$

where

- l: Zipcodes; t: month;
- Q_{lt}: total quantity (total size or total number of applications); P_{lt}^{MKT}: average price ($ per watt) of a solar panel;
- I_{ltd}: indicator of one month before a step change;
- X_{lt} : Other local control variables, such as house price, solar-radiation intensity, and political view.

Therefore:

$$
\hat{\gamma} = \frac{\hat{\beta}^q}{\hat{\beta}^p}; \quad \hat{\delta} = \frac{\hat{\beta}^q \hat{\gamma}}{\hat{\beta}^q - \hat{\gamma}}.
$$

Table 1.7 presents results using only Zipcode-months with at least one application. Signs for all coefficients and estimators are consistent with my theory. β_c and β_{size} indicate the impact of rebates on quantity. As rebates increase, equilibrium quantity should increase. Positive β_p represents that as rebates increase, the market price of solar panels would increase. δ and γ represent estimated demand and supply coefficients. Negative δ means that as the price increases for households, demand would decrease, while positive γ means that the price increases for sellers, supply would increase. I find that as the price of solar panels increases by $1 per watt, 0.8 to 1.5 fewer households would demand solar panels, which is related to a 6- to 9-watt decrease in solar capacity. On the other hand, as the price of solar panels increases by $1, sellers are willing to sell 0.5 to 0.6 more solar panel systems and 3.5 to 4.3

watts more capacity. These results are consistent across number of months before and after a CSI-step switch.

Table 1.7: Results Using Only Positive Application Zipcode-Month

	All		Months before and after step switch					
			6 months		5 months		4 months	
	(1)		(2)		(3)		(4)	
	Coef.	std.err	Coef.	std.err	Coef.	std.err	Coef.	std.err
β_c	.0710	.087	.413***	.091	.409***	.093	.414***	.101
β_{size}	.5253	.592	2.51***	.616	2.85***	.633	2.94***	.694
β_p	.734***	.049	.724***	.056	.675***	.060	.676***	.063
δ_c	–.2671	.353	–1.50***	.516	–1.26***	.427	–1.28***	.464
γ_c	.097	.119	.570***	.135	.605***	.149	.613***	.162
δ_{size}	–1.976	2.37	–9.12***	3.21	–8.77***	2.79	–9.05***	3.09
γ_{size}	.715	.816	3.470***	.9187	4.22***	1.04	4.35***	1.15
N	33,340		29,026		27,526		25,294	

			3 months		2 months			
			(5)		(6)			
			Coef.	std.err	Coef.	std.err		
β_c			.314***	.109	.037	.134		
β_{size}			2.40***	.719	.721	.929		
β_p			.610***	.072	.409***	.089		
δ_c			–.805**	.361	–.062	.238		
γ_c			.514***	.189	.090	.351		
δ_{size}			–6.15***	2.38	–1.22	1.64		
γ_{size}			3.92***	1.30	1.76	2.54		
N			21,277		14,887			

β_c: coefficient of rebate on number of applications
β_{size} : coefficient of rebate on total size(in kW) of applications
β_p : coefficient of rebate on per unit (watt) price of solar panel
δ_c : demand estimator of solar when the quantity is measured in quantity
γ_c : supply estimator of solar when the quantity is measured in quantity
δ_{size} : demand estimator of solar when the quantity is measured in total size
γ_{size} : supply estimator of solar when the quantity is measured in total size

Column (1) displays the results using all the data. The coefficients and estimators are not significant. This is because the demand and supply shock is not well controlled when considering a long period before and after the step change. Columns (2) to (6) display the results using only a certain month before and after the CSI-step change. The coefficients and estimators are statistically significant, and the level doesn't change much as the number of months changes.

In all the regression specifications, I include an indicator for one month before the step change, because households may be more eager to apply for the rebate at the end of a step, before the rebate rate decreases, and contractors may try harder to incentivize households to do so at the end of the step as well. Including an indicator for one month before the step change can control for this issue, but this makes it harder for me to identify any quantity and price change due to rebate level during a short period before and after the step change. I present data only up to six months before and after the step change, since in Table 1.3, the average gap between steps is around eight to nine months.

1.5.3 Elasticity

Earlier in this section, I presented results of the impact of price on quantity for demand and supply under different econometric models. However, it would be hard to compare them without unifying the unit of the impact. I calculate the elasticity for demand and supply for these two model as follow:

- **Define demand and supply elasticity**

$$\epsilon^D = \frac{\partial E(Q^D)}{\partial Price^{cons}} \frac{Price^{cons}}{E(Q^D)} \tag{1.35}$$

$$\epsilon^D = \frac{\partial E(Q^S)}{\partial Price^{mkt}} \frac{Price^{mkt}}{E(Q^S)}. \tag{1.36}$$

- **Two-part model by Duan et al. (1983)**

For the two-part model, the overall expected quantity is

$$\begin{aligned}
\text{Demand:}\quad E(q_{it}|Price_{it}^{cons}, X_{it}) =&(1-\Phi(\alpha_1^D+\delta_1 Price_{it}^{cons}+\beta_1^D X_{1it}))\\
&\times exp(\frac{\sigma_{2D}^2}{2})\times exp(\alpha_2^D+\delta_2 Price_{it}^{cons}+\beta_2^D X_{2it})\\
\text{Supply:}\quad E(q_{it}|Price_{it}^{mkt}, X_{it}) =&(1-\Phi(\alpha_1^S+\gamma_1 Price_{it}^{mkt}+\beta_1^S X_{1it}))\\
&\times exp(\frac{\sigma_{2S}^2}{2})\times exp(\alpha_2^S+\gamma_2 Price_{it}^{mkt}+\beta_2^S X_{2it}).
\end{aligned}$$

Based on the above quantity specification, I calculated demand and supply elasticity to be

$$\epsilon^D = \Big[\frac{-\phi(\alpha_1^D+\delta_1 Price_{it}^{cons}+\beta_1^D X_{1it})}{(1-\Phi(\alpha_1^D+\delta_1 Price_{it}^{cons}+\beta_1^D X_{1it}))}\times\delta_1+\delta_2\Big]Price^{cons} \tag{1.37}$$

$$\epsilon^S = \Big[\frac{-\phi(\alpha_1^S+\gamma_1 Price_{it}^{mkt}+\beta_1^S X_{1it})}{(1-\Phi(\alpha_1^S+\gamma_1 Price_{it}^{mkt}+\beta_1^S X_{1it}))}\times\gamma_1+\gamma_2\Big]Price^{mkt}. \tag{1.38}$$

- **Using linear regression to model only Zipcode-month with positive applications**

$$\epsilon^D = \delta\frac{Price^{cons}}{Q^D};\quad \epsilon^S = \delta\frac{Price^{mkt}}{Q^S} \tag{1.39}$$

Table 1.8 includes the results for an elasticity comparison across my linear model using only Zipcode-months with at least one application and the two-part model including zeros. I use sample average price, quantity, and probit probability to estimate elasticity for these models. For the linear model Zipcode-months with at least one application, I find the demand elasticity to be –1.2 to –1.5 if all months are being used. This is inline with other literature that studies the elasticity of solar-panel demand. Gillingham and Tsvetanov (2016) found demand elasticity to be –1.76 using data from Connecticut. Focusing on a certain period before and after the CSI-step switch greatly affects the results. Demand elasticity becomes –4 to –7.5, and supply elasticity increases to 3 to 4. Both demand and supply elasticity become more inelastic when controlling for a shorter period, because it is harder to make

Table 1.8: Elasticity

	All	6 Months	5 Months	4 Months	3 Months	2 Months
			Linear Using Only Positive Zipcode-Months			
ϵ_C^D	–1.193*	–7.002***	–5.995***	–6.121***	–3.888***	–0.310
	(0.695)	(1.133)	(0.894)	(0.990)	(0.755)	(0.509)
ϵ_C^S	0.514*	3.159***	3.413***	3.474***	2.927***	0.528
	(0.268)	(0.319)	(0.362)	(0.404)	(0.479)	(0.867)
ϵ_{size}^D	–1.507*	–7.371***	–7.253***	–7.535***	–5.157***	–1.043*
	(0.794)	(1.243)	(1.013)	(1.155)	(0.867)	(0.603)
ϵ_{size}^S	0.649**	3.325***	4.129***	4.277***	3.882***	1.778*
	(0.314)	(0.381)	(0.441)	(0.486)	(0.571)	(0.999)
			Two-Part Model			
Use Price						
ϵ^D	–1.468***	–2.565***	–2.789***	–2.958***	–3.284***	–2.74***
	(–0.314)	0.262	0.305	0.301	0.327	0.331
ϵ^S	1.451***	3.174***	3.664***	4.173***	5.572***	8.372**
	(0.447)	0.547	0.693	0.725	0.997	3.474
Use ln(Price)						
ϵ^D	–4.583***	–5.73***	–5.99***	–6.195***	–6.389***	–5.522***
	(0.302)	(0.344)	(0.335)	(0.363)	(0.338)	(0.358)
ϵ^S	5.921***	11.661***	12.369***	15.713***	21.781***	47.101
	(0.555)	(1.453)	(1.505)	(2.562)	(4.214)	(110.697)

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

changes in a short amount of time. It may take some time for the market to respond. My results are consistent when using number of applications and total size of applications.

The results using the two-part model shows a larger variance between using Price and $ln(Price)$. Similar to the linear model, when focusing on the period around a CSI-step change, estimated elasticities are higher. It is easier to control for unobserved shifts. Using price as the control, demand elasticity is –1.4 using all data. This is almost the same as my linear model estimations. Looking at periods around CSI-step changes, demand elasticity is around –3. This is less elastic compared with the estimation from my linear model. However, supply elasticity is higher for both using all data and using a few months before and after a step change. Supply elasticity is between 3 and 8. When using $ln(Price)$ as the price, estimated elasticities are higher in absolute value for both demand and supply. Compared with elasticities from my linear model, demand elasticity is closer but still smaller in absolute value, while supply elasticity is much bigger. Estimated demand elasticity is around –6, while estimated supply elasticity is between 11 and 22.

From this table, I can conclude that including zero greatly affects estimated elasticities and leads to different conclusions. Using only positive application data and my linear model, one would conclude that demand is more elastic than supply; therefore, a rebate would increase a seller's profit more than it would increase consumer surplus. However, using my two-part model to analyze the data, including markets with zero quantity, the conclusion is the opposite. I would argue that my two-part model is a better model, since it accounts for impact of price on zero distribution.

1.6 Heterogeneous Effect

1.6.1 County Elasticity Estimation

California is the most populous U.S. state, and the third largest by area. Therefore, preference for rooftop solar-panel systems varies substantially across California. In this section, I estimate county-specific demand and supply elasticities. This information would help policy makers target areas with the highest rebate efficiency. I use the following two-part model to estimate county-specific elasticities:

- Probit:

$$Pr(d_{it} = 0) = \Phi(\alpha_1 + \sum_c \theta_{1c} I(c) Price_{it} + \beta_1 X_{it}) \tag{1.40}$$

- Log-normal for positive quantity:

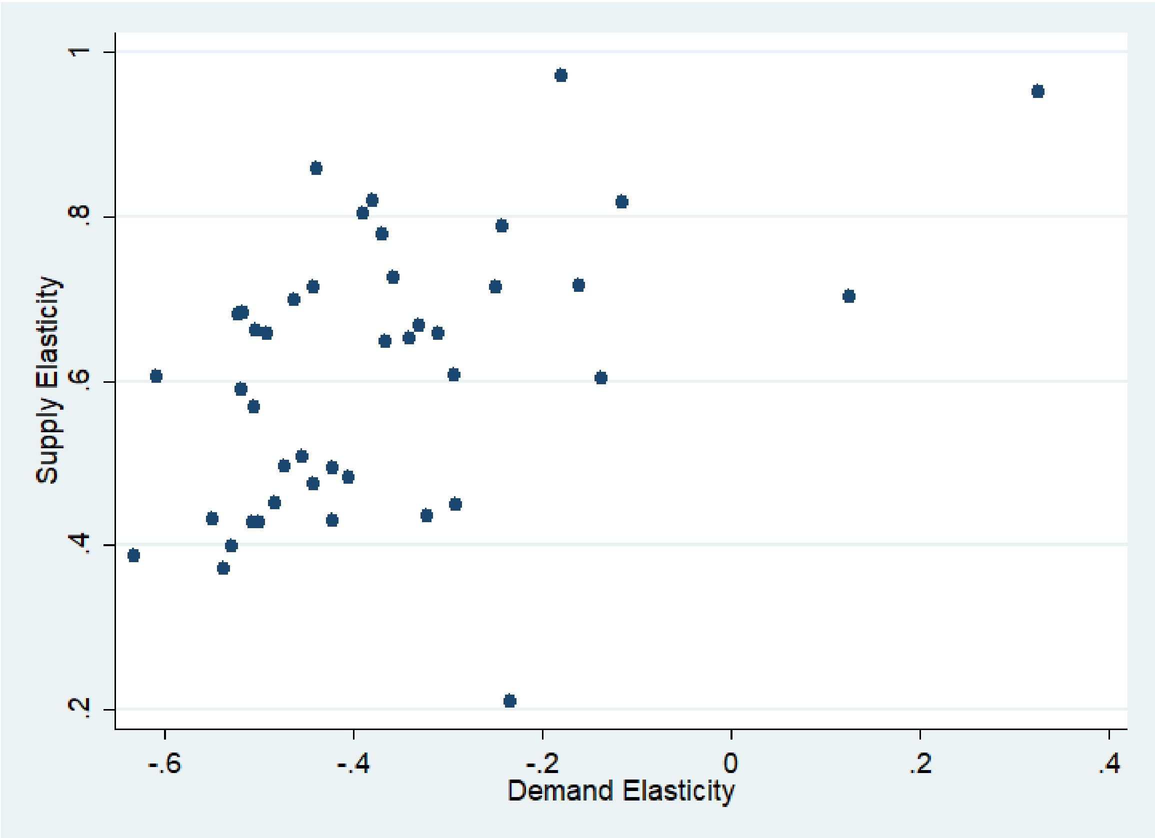
$$log(q_{it}|q_{it} > 0) = \alpha_2 + \sum_c \theta_{2c} I(c) Price_{it} + \beta_1 X_{it}, \tag{1.41}$$

where *price* is $Price^{cons}$ for the demand estimation and $Price^{mkt}$ for the supply estimation. c represents county. $I(c)$ is an indicator for each county, $\theta = \delta$ for demand estimation, and $\theta = \gamma$ for supply estimation. I estimate this model using control function illustrated in

Section 1.3. For elasticity estimation, I use the function derived in Section 1.5:

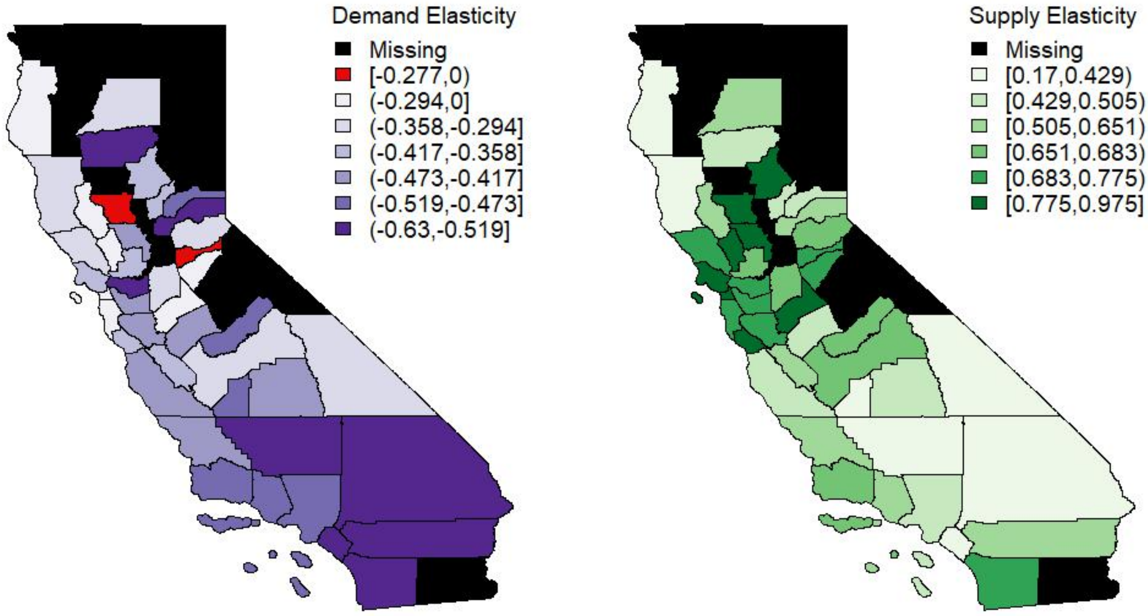
$$\epsilon_c = \Big[\frac{-\phi(\alpha_1 + \theta_{1c}Price_{it} + \beta_1 X_{1it})}{(1 - \Phi(\alpha_1 + \theta_{1c}Price_{it} + \beta_1 X_{1it}))} \times \theta_{1c} + \theta_{2c}\Big]Price.$$

Figure 1.10: County Elasticity



The results for elasticity are shown in Figure 1.10. In this figure, each point represents a county in California. All counties' estimated supply elasticities are positive, but only two counties have positive demand elasticities. Supply elasticities range from 0.2 to 1. This small range indicates that all counties studied have inelastic supply. Given that it's negative, demand elasticity ranges between zero to –0.7. This also implies that all counties have inelastic demand. There seems to be a negative relationship between demand and supply elasticities. This means higher demand elasticity is correlated with lower supply elasticities. Figures 1.11(a) and 1.11(b) put the elasticity estimation onto a map of California. The black

Figure 1.11: County Level Elasticity



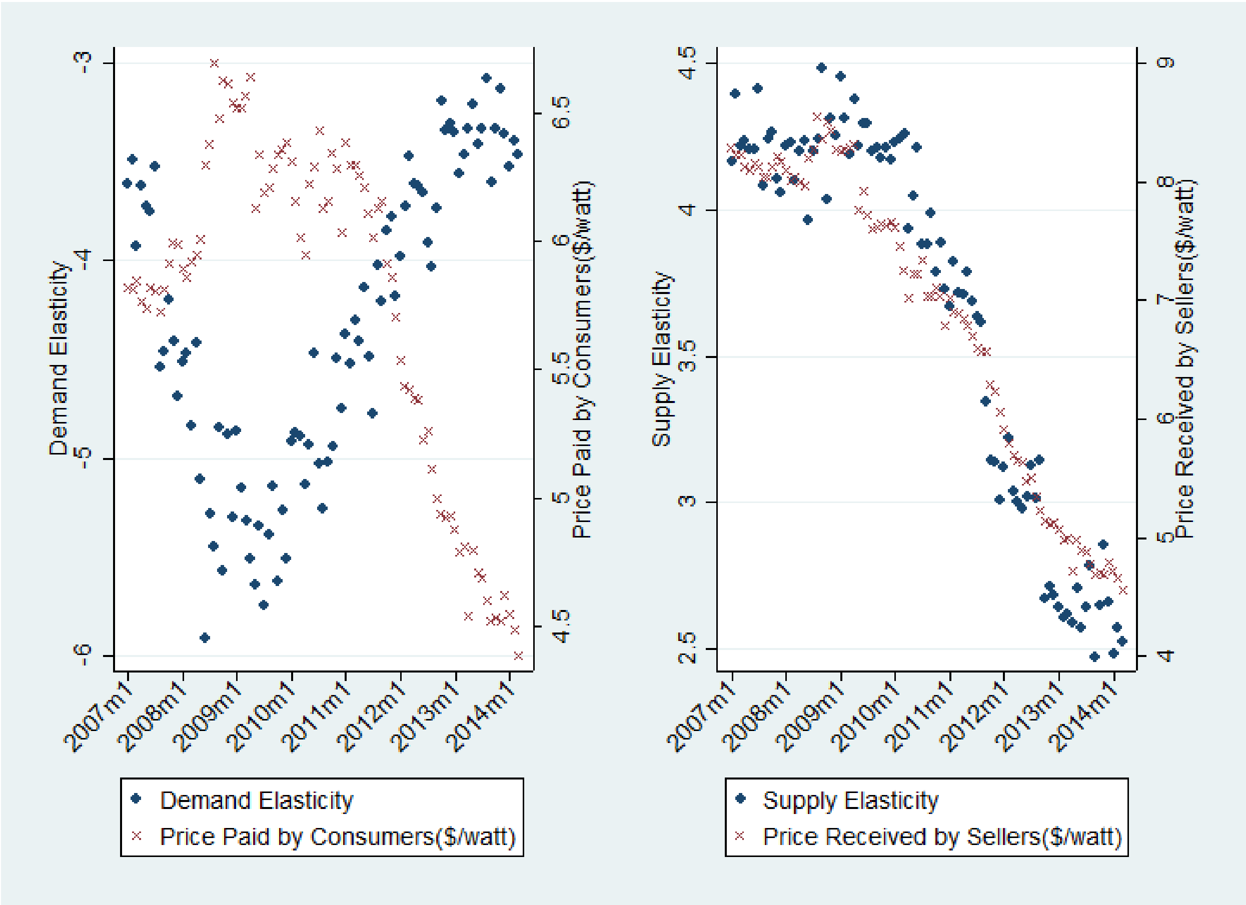
(a) County-Level Demand Elasticity Map (b) County-Level Supply Elasticity Map

area represents the ones that did not participate in the CSI. In the demand elasticity map, the two red counties are the ones with positive demand elasticity. These two counties are Amador and Colusa,two of the least populated counties in the state. These two figures show some counties with higher demand but lower supply elasticity, such as San Bernardino and Kern. There are also counties with low supply and demand elasticities, such as Coastal and Northern California. The Bay Area has low demand elasticity but slightly higher supply elasticity. Based on these two figures, if policy designers wanted to transfer rebates to consumers more, they should target the Bay Area; if they wanted to help sellers more, they should target Southern California. Coastal and Northern California are areas where rebates would not be very effective in increasing quantity.

1.6.2 Elasticity Over Time

In this section, I study the elasticity over time. This will help us to understand the trend of elasticity change and use it to predict the future and provide better policy guidance.

Figure 1.12: Elasticity Against Time



In Section 1.5, I estimate elasticity using average price and quantity. However, I understand that the difference between elasticity and price is important. Figure 1.12 presents the relationship between estimated elasticity at the average price for each month. I observe a strong positive relationship both between consumer price and demand elasticity and between market price and supply elasticity. This indicates that when the price is high, consumers and sellers alike are highly sensitive to price. When the price of solar panels is low, consumers and sellers alike are less sensitive to price changes. This indicates that if the price were to continue to decrease in the future, the market will be less and less sensitive to the price change of solar panels. This may imply that future rebate programs would not to be as effective.

The above estimation uses the same demand and supply function, but it estimates elas-

Table 1.9: Elasticity by Step Change

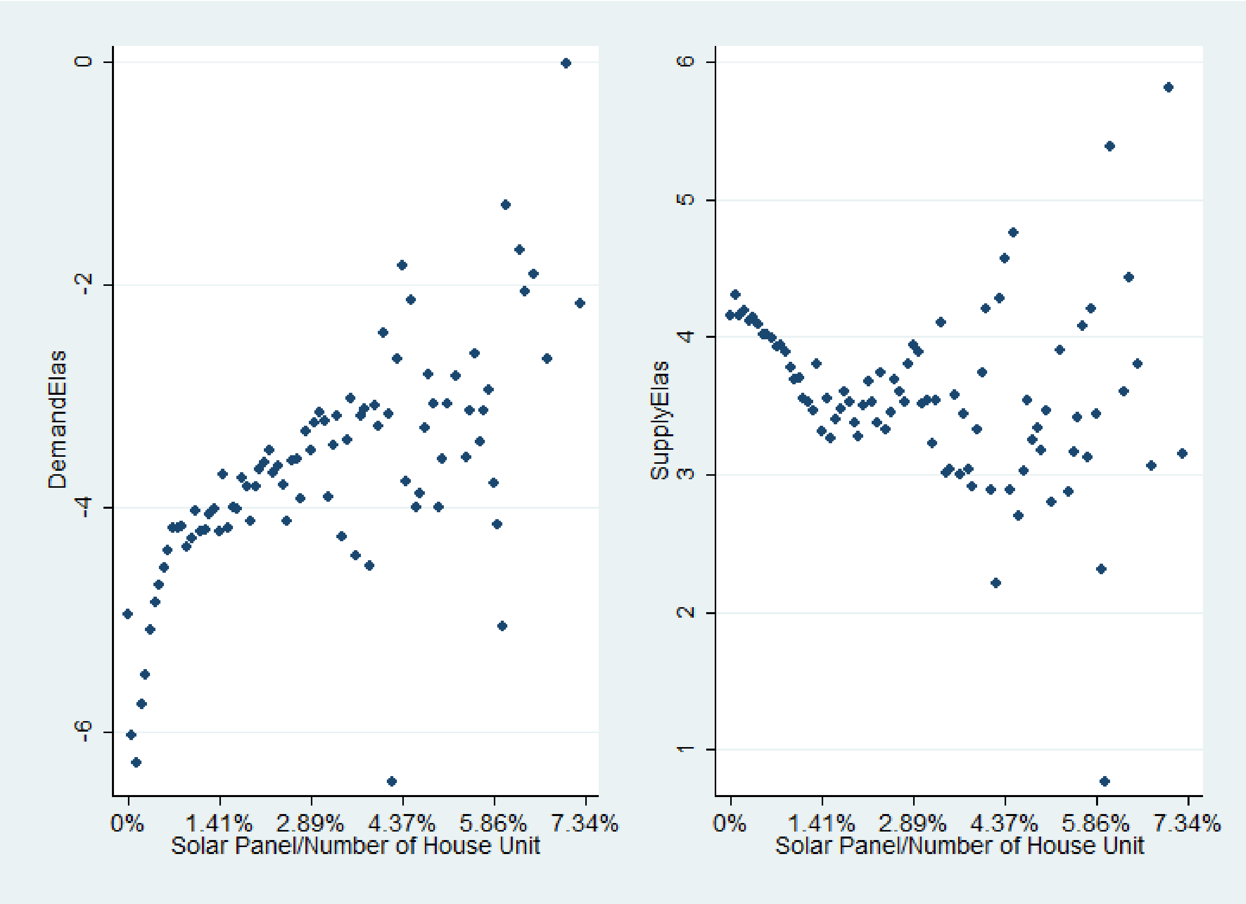
	Elasticity		Demand Coef.		Supply Coef.	
Step	ϵ^D	ϵ^S	δ_1	δ_2	γ_1	γ_2
Step 3	–6.98***	10.89***	1.053***	–0.105	–1.044***	0.463
	(0.629)	(3.852)	(0.013)	(0.102)	(0.012)	(0.468)
Step 4	–6.616***	26.82	0.780***	–0.234***	–0.743***	2.840
	(0.621)	(34.116)	(0.010)	(0.097)	(0.221)	(7.757)
Step 5	–6.088***	9.462	0.700***	–0.219**	–0.512	0.999
	(0.649)	(69.540)	(0.007)	(0.103)	(0.477)	(9.500)
Step 6	–7.273***	–2575	0.700***	–0.428***	0.151	–355.211
	(0.924)	(25030)	(0.008)	(0.150)	(0.685)	(3344.433)
Step 7	–7.129***	–18.87	0.772***	–0.365	0.365	5.822
	(2.151)	(114.081)	(0.009)	(0.350)	(0.684)	(65.503)
Step 8	–2.672	–2.515	0.955***	0.519	–0.704	–1.914
	(8.514)	(20.953)	(0.009)	(1.416)	(0.648)	(9.608)
Step 9	–5.953*	3.007	1.164***	0.103	0.467	5.498
	(3.613)	(50.066)	(0.011)	(0.646)	(1.071)	(38.754)

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

ticity with different price and quantity levels. To further check whether the demand and supply functions for rooftop solar panels is changing over time, I estimate demand and supply elasticity at each point there is a CSI-step change. Table 1.9 presents the results. These results show that demand elasticity is quite consistent and does not change much over time. However, the supply elasticity varies greatly and is not significant for most of the step changes. This can be due to my small sample when I look only at a small period before and after a single CSI-step change.

I further estimate how elasticity changes as the ratio of households with solar panels in the given Zipcodes rises. The result is presented in Figure 1.13. Note that the highest percentage of solar panel system to the number of unit housing is 7%. This shows that the solar panel market still has a lot of potential. From this figure, I observe a strong negative correlation between demand elasticity and the percentage of households with solar panels. This indicates that the first households that decide to have solar panels are the ones with high demand elasticity. This further emphasizes the previous result, that demand elasticity is high when price level is high. As more and more households in the Zipcodes install solar

Figure 1.13: Elasticity over Previous Ratio of Households with Solar Panel Systems in the Zipcodes



panels, the price goes down and demand becomes more inelastic. This is because people are less sensitive to price changes when the price level is low. Supply elasticity follows a slightly different pattern. When there are few installations in the area, elasticity decreases as there are more households installing solar panels. However, the variance increases as the percentage of households with solar panels increases.

1.7 Welfare Analysis

1.7.1 Surplus and Deadweight Loss

Knowing the elasticity level would allow me to estimate changes in consumer surplus, producer surplus, and deadweight loss generated by the CSI program. Given a demand and supply model, I estimate surplus follows:

$$\Delta CS = \int_{P^{cons}}^{P^{Eq}} E(Q^D|P)dP$$
$$\Delta PS = \int_{P^{Eq}}^{P^{mkt}} E(Q^S|P)dP$$
$$DWL = E(Q) * Rebate - (\Delta CS + \Delta PS),$$

where ΔCS and ΔPS are changes in the consumer surplus and producer surplus. P^{cons} and P^{mkt} are, respectively, the price paid by households and the price received by producers. P^{Eq} is the equilibrium price, which is calculated using the impact of rebates on prices. $E(Q^D|P)$ and $E(Q^S|P)$ are the demand and supply functions given a certain price level.

Table 1.10: Surplus Generated by CSI

	Income Quantile				
	Q1	Q2	Q3	Q4	Total
Consumer surplus	17,455	55,894	79,220	116,554	269,123
	(6.49%)	(20.77%)	(29.44%)	(43.31%)	
Producer surplus	17,334	54,742	75,262	109,203	256,540
	(6.76%)	(21.34%)	(29.34%)	(42.57%)	
Total surplus	34,789	110,636	154,482	225,757	525,663
	(6.62%)	(21.05%)	(29.39%)	(42.95%)	
Deadweight loss	44,412	139,768	195,008	284,230	663,417
	(6.69%)	(21.07%)	(29.39%)	(42.84%)	
Total surplus / Total cost	43.92%	44.18%	44.20%	44.27%	44.21%

Table 1.10 presents the estimated change in surplus caused by rebates using supply and demand estimation with all data. Since the demand and supply have similar elasticity when estimated using all data, consumer surplus and producer surplus have similar changes. Consumer surplus increases by \$269 million and producer surplus increases by \$256 million due to CSI. However, CSI also generated a total of \$663 million in deadweight loss, which is the amount of expenditure that increases neither consumer nor producer surplus. I call this deadweight loss in this case because I consider the impact of rebate only on consumer and producer surplus, ignoring the externality generated by solar panels on the environment, which I don't cover here. In this paper, I estimate that only 44% of money spent in the CSI

increases either consumer or producer surplus.

I further study the welfare changes across income quantiles. Both consumer and producer surplus increase more with higher levels of income. For the bottom 25% of income Zipcodes, consumer total surplus increased by less than $34 million, less than 7% of the total surplus increase. The top 25% of income Zipcodes got the highest surplus increase, more than 40% ($226 million). This is because high-income Zipcodes buy more solar panels earlier, when the rebate level is high.

1.7.2 Impact of the Rebate

This section provides additional analysis of the impact of CSI rebate on quantity and price. Most of the result comes directly from Section 1.5, except for the impact of the rebate on quantity, including zero-quantity distribution. I use the following specification to study the rebate's impact on quantity:

$$\begin{aligned}\frac{\partial E(q_i)}{\partial Rebate_i} &= \frac{\partial E(q_i)}{\partial Price_i}\frac{\partial Price_i}{\partial Rebate_i} \\ &= \Big[\frac{-\phi(\alpha_1 + \theta_1 Price_{it} + \beta_1 X_{1it})}{(1-\Phi(\alpha_1 + \theta_1 Price_{it} + \beta_1 X_{1it}))} \times \theta_1 + \theta_2\Big] E(q_i)\lambda,\end{aligned}$$

where σ^2 is the variance of the error term in fitting positive quantity with log-normal. θ_1 and θ_2 are coefficients of the impact of price on quantity. θ_1 is the coefficient on predicting zero and θ_2 is the coefficient on fitting positive quantity. λ represents the impact of the rebate on price.

Table 1.11 organizes the results of the rebate's impact on quantity and price, using positive-only quantity and including zero quantity. The results are quite different from these two specifications. One dollar of rebate on solar panels increases the average number of households that install solar panels at a given Zipcode-month by around 0.4 if only positive quantity is considered. However, when I account for the change in zero distribution, the expected number of household install solar panels increases to 1.4. Based on number of markets and their rebate rate, I estimate that the CSI led to 33,864 more households having

Table 1.11: Impact of Rebate on Quantity and Price

	All	6 Months	5 Months	4 Months	3 Months	2 Months
			Impact of Rebate on Quantity			
Linear model on Positive Quantity						
β_c	0.0710	0.413***	0.409***	0.414***	0.314***	0.037
	(0.087)	(0.091)	(0.093)	(0.101)	(0.109)	(0.134)
β_{size}	0.5253	2.51***	2.85***	2.94***	2.40***	0.721
	(0.592)	(0.616)	(0.633)	(0.694)	(0.719)	(0.929)
Two-Part Model Including Zero Quantity						
β_c	0.554***	0.781***	0.816***	0.899***	1.411***	18.566
	(0.052)	(0.085)	(0.102)	(0.136)	(0.416)	(59.831)
β_{size}	2.735***	3.816***	4.160***	4.615***	7.512***	1687.654
	(0.294)	(0.551)	(0.665)	(0.908)	(2.154)	(14097)
			Impact of Rebate on Market Price of Rebate			
Linear model on Positive Quantity						
β_p^{mkt}	0.734***	0.724***	0.675***	0.676***	0.610***	0.409***
	0.049	0.056	0.060	0.063	0.072	0.089
Two-Part Model including Zero Quantity						
β_p^{mkt}	0.455***	0.360***	0.356***	0.337***	0.287***	0.146***
	(0.021)	(0.025)	(0.026)	(0.027)	(0.029)	(0.036)

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

solar panels if not considering zero, but 119,456 more households if I include Zipcodes with zero applications. The estimated impact of the rebate on total capacity also depends on whether the specification includes Zipcodes with zero applications. My linear model using Zipcodes with nonzero quantity led me to estimate that one dollar in rebate increases capacity by 2.4 to 2.9 kW per Zipcode-month, and 2.7 to 7.5 kW using my two-part model including Zipcodes with zero applications. Using the total rebate amount, I estimate that total increased capacity due to the CSI rebate is between 203 and 253 MW without zero and between 228 and 634 MW if I include zero.

The impact of the rebate on prices also differs across specifications. One dollar in rebate increases the price received by sellers by around \$0.7 with only positive quantity , and around \$0.35 with zero quantity. Assume that my estimation of the impact of price on distribution of Zipcode-months with zero quantity is accurate. This difference in estimation indicates that when the rebate decreases, market with high prices are more likely to have zero quantity.

Therefore, when do analyzing only positive-quantity markets, I observe a large change in market price. The above estimations also imply a passthrough rate of rebate to households and sellers. Using positive quantity, I estimate that for each dollar of rebate, around 70% goes to sellers to increase their selling price, while 30% goes to households to decrease the price they pay. Including zero distribution, around 35% goes to sellers while 65% goes to households.

1.8 Robustness Check

In this section, I test another zero-inflated model: zero-inflated negative binomial(ZINB). However, the ZINB is more limited than the two-part model because it focuses on count data. Yet, household decisions about solar panels are not only whether to have one but also the size of the system, which is a continuous variable.

1.8.1 Zero-Inflated Negative Binomial(ZINB) Model

Here, I show the ZINB model, which can be used to fit excess zero and count data and compare the results with the two-part model and the linear using positive-only data.

Statistical model

- Probability Function:

$$Pr(y_{ct} = j) = \begin{cases} \pi_{ct} + (1 - \pi_{ct})g(y_{ct} = 0) & \text{if } j = 0 \\ (1 - \pi_{ct})g(y_{ct}) & \text{if } j > 0. \end{cases} \tag{1.42}$$

- π_{ct} represents the probability that a household in a Zipcodes in a given month is not thinking about buying solar.

  $$\pi_{ct} = \frac{\lambda_{ct}}{1 + \lambda_{ct}},$$

  where

  $$\lambda_{ct} = exp(\zeta_1 z_{1ct} + \zeta_2 z_{2ct} + ... + \zeta_m z_{mct}).$$

- $g(y_{ct})$ follows a negative binomial distribution:

$$g(y_{ct}) = Pr(Y = y_{ct}|\mu_{ct}, \alpha) \tag{1.43}$$

$$= \frac{\Gamma(y_{ct} + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_{ct} + 1)} \Big(\frac{1}{1 + \alpha\mu_{ct}}\Big)^{\alpha^{-1}} \Big(\frac{\alpha\mu_{ct}}{1 + \alpha\mu_{ct}}\Big)^{y_{ct}}, \tag{1.44}$$

  where

$$E[y_{ct}|x_{ct}] = \mu_{ct} = exp(\beta_1 x_{1ct} + \beta_2 x_{2ct} + ... + \beta_k x_{kct}).$$

I would like to use this model to identify the impact of price on quantity. Since the expected value of quantity is a linear function of controls, I can use the method I developed for linear demand and supply. This means I estimate the impact of rebates on quantity and on price, then I back out the impact of price on quantity. I first set up the demand and supply model to be as follows: Demand:

$$E(Q^d) = exp\{\alpha^d + \delta P^{Cons} + \theta^d X + \epsilon^d\}. \tag{1.45}$$

Supply:

$$E(Q^s) = exp\{\alpha^s + \gamma P^{MKT} + \theta^s X + \epsilon^s\}. \tag{1.46}$$

The equilibrium condition would be:

$$E(Q^d) = E(Q^s)$$

$$P^{Cons} + Rebate = P^{Mkt}.$$

I use equilibrium to solve the above functions:

$$E(Q) = exp\{\alpha^q + \frac{\delta\gamma}{\delta - \gamma} Rebate + \theta^q X + \epsilon^q\} \tag{1.47}$$

$$lnP^{MKT} = \alpha^p + \frac{\delta}{\delta - \gamma} Rebate + \theta^p X + \epsilon^p. \tag{1.48}$$

Specification

To estimate this model, I use the following specification. I first use the ZINB model to fit quantity:

$$E(Q_{lt}) = exp\{\alpha^q + \beta^q Rebate_{lt} + \theta^q X_{cy} + \zeta_t^q + \sum_{d=1}^{K} \iota_d^q I_{ltd} + \epsilon_{lt}^q\}. \quad (1.49)$$

I use the linear model to fit price:

$$P_{lt}^{MKT} = \alpha^p + \beta^p Rebate_{lt} + \theta^p X_{cy} + \zeta_t^p + \sum_{d=1}^{K} \iota_d^p I_{ltd} + \epsilon_{lt}^p, \quad (1.50)$$

where

- l: Zipcodes-utility; t: month;
- Q_{lt}: quantity (either total size or total number of applications); P_{lt}^{MKT}: average price ($ per watt) of a solar panel;
- I_{ltd}: indicator of one month before a CSI-step change.

Therefore,

$$\hat{\gamma} = \frac{\hat{\beta}^q}{\hat{\beta}^p}; \quad \hat{\delta} = \frac{\hat{\beta}^q \hat{\gamma}}{\hat{\beta}^q - \hat{\gamma}}$$

Demand and supply elasticity estimations are according to the following function:

$$\epsilon^D = \delta Price^{cons}; \quad \epsilon^S = \gamma Price^{mkt}. \quad (1.51)$$

Table 1.12 contains the results of my quantity and price regressions. I measure the quantity using either number of applications or total size (kW) of applications. I modeled quantity regressions using the ZINB model. Both quantity regressions have a positive coefficient of rebate. This represents a strong positive correlation between rebate and quantity when accounting for zero quantity. However, the coefficient of rebate on total size of applications is not statistically significant. The effect of rebate on price is positive. It means that, as the rebate increases by $1, the market price of solar panels would increase by $0.723. Thus, the rebate passthrough rate to sellers is around 72% and to households is around 27%.

Table 1.12: Zero-Inflated Negative Binomial model

VARIABLES	Count		Total Size		$price^{MKT}$	
	Coef.	Std.Err	Coef.	Std.Err	Coef.	Std.Err
Rebate	0.186***	(0.017)	0.017	(0.016)	0.723***	(0.016)
Previous count	0.019***	(0.000)	0.020***	(0.000)	–0.001***	(0.000)
I(one month before step change)	0.346***	(0.013)	0.345***	(0.012)	–0.251***	(0.017)
Total voter registration	0.261***	(0.006)	0.054***	(0.006)	0.131***	(0.005)
Average house price(1,000,000)	0.368***	(0.031)	0.744***	(0.030)	–0.179***	(0.028)
House price per Sq.f.t(1000)	–0.343***	(0.075)	–1.606***	(0.072)	0.947***	(0.067)
Solar intensity	0.915***	(0.074)	0.475***	(0.067)	2.016***	(0.363)
Ratio of Democrats	–0.967***	(0.048)	–0.979***	(0.045)	0.482***	(0.041)
Constant	–7.391***	(0.327)	–0.745**	(0.298)	5.265***	(0.063)
Observations	68,701		68,701		68,670	
Inflate						
Rebate	120.140***	(2.155)	2.322***	(0.032)		
Previous count	–0.012***	(0.002)	–0.051***	(0.001)		
I(one month before step change)	–16.595***	(0.530)	–0.844***	(0.032)		
Total voter registration	–0.410***	(0.046)	–0.392***	(0.012)		
Solar intensity	–19.664***	(3.413)	0.334**	(0.150)		
Ratio of Democrats	3.202***	(0.315)	2.052***	(0.082)		
Constant	–277.394	(1,890.044)	–40.989***	(0.673)		
Observations			68,701			

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To better control for unobserved market shifters, Table 1.13 shows the results using a certain month before and after a CSI-step switch. The estimator of the impact of rebates on number of applications is consistent across months, between 0.11 and 0.25. The effect of rebates on the market price of solar panels is between 0.77 and 0.93. This is slightly higher than the estimators using only Zipcode-month with at least one application. Unfortunately, demand and supply estimators have large variance in this setup.

Table 1.13 also presents the elasticity estimation. However, other than supply elasticity using count as quantity, other estimations have the sign I expect but are not statistically significant. The results using the ZINB model indicate that it is inaccurate when used

Table 1.13: Zero-Inflated Negative Binomial model by Months from Step Switch

	All Months	6 Months	5 Months	4 Months	3 Months	2 Months
β_c	0.028	0.111***	.133***	0.166***	0.251***	0.170***
	(0.019)	(0.019)	(.020)	(0.021)	(0.024)	(0.030)
β_{size}	0.016	0.061***	.026	0.021	–0.011	–0.046*
	(0.020)	(0.020)	(.022)	(0.024)	(0.020)	(0.026)
β_p	0.949***	0.931***	.900***	0.910***	0.858***	0.769***
	(0.040)	(0.042)	(.043)	(0.045)	(0.053)	(0.065)
δ_c	–0.551	–1.619	–1.326	–1.842	–1.770	–0.737**
	(121.227)	(34.17)	(31.99)	(30.62)	(4.780)	(0.314)
γ_c	0.029	0.119***	.148***	0.182***	0.293***	0.221***
	(0.022)	(0.022)	(.023)	(0.024)	(0.029)	(0.043)
δ_{size}	–0.318	–0.885	–.257	–0.233	0.076	0.201
	(5.011)	(14.83)	(12.90)	(5.046)	(0.881)	(0.161)
γ_{size}	0.017	0.065***	.029	0.023	–0.013	–0.060*
	(0.022)	(0.022)	(.024)	(0.026)	(0.024)	(0.034)
			Elasticity Estimation			
ϵ_C^D	–3.209	–9.560	–7.885	–10.95	–10.50	–4.362**
	(706.0)	(203.3)	(190.8)	(183.0)	(28.48)	(1.867)
ϵ_C^S	0.203	0.837***	1.041***	1.282***	2.048***	1.544***
	(0.147)	(0.147)	(0.157)	(0.164)	(0.196)	(0.292)
ϵ_{size}^D	–1.854	–5.256	–1.528	–1.386	0.454	1.190
	(29.18)	(88.51)	(76.92)	(30.16)	(5.248)	(0.955)
ϵ_{size}^S	0.117	0.460***	0.202	0.162	–0.089	–0.421*
	(0.142)	(0.151)	(0.164)	(0.177)	(0.161)	(0.229)

Standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

to model continuous distribution. The ZINB model performs better when the dependent variable is number of applications, rather than total size. However, even with number of applications, it failed to capture the demand and supply function. I believe this is because the market is dependent on not only how many household want to buy solar panel but also how much each household would like to buy. Therefore, considering only the number of households or number of applications would make it hard to capture the true demand and supply pattern.

1.9 Conclusion

In this paper, I study demand and supply elasticity of rooftop solar-panel systems using the California Solar Initiative, a capacity-based rebate program. This program provides a special setup—the exogeneous change in rebate rate—that allows me to study both demand and supply at the same time. I contribute to the line of research in two ways. First, I develop a new way of using the rebate rate shift and its impact on price and quantity to derive both demand and supply functions. Second, I analyze the disaggregated market at the Zipcode-month level. I can use local variations to control shifts in demand and supply, though this creates a problem: that a large portion of the market would have zero solar-panel rebate applications. To overcome this problem, I use a two-part model to incorporate Zipcode-months with zero solar applications. The result indicates that considering zero quantity in the market is crucial for correct estimation.

I estimate that a 1% increase in price would decrease the demand for rooftop solar-panel capacity 3.3% and increase the supply of rooftop solar panels 5.6%. This estimation comes from my two-part model incorporating Zipcode-months with zero solar-panel applications and focuses on the three months before and after a rebate rate change. This two-part model is better than a linear model that uses only Zipcode-months with positive solar-panel applications, because it considers the change in distribution of zero caused by a change in price.

I use elasticity estimations to derive the impact of the CSI and to conduct welfare analysis. I find that the CSI caused a large increase in residential solar-panel adoption and solar-power-generating capacity. It also benefited households more than sellers. However, the CSI unfortunately generated a huge deadweight loss, creating inefficiency. Moreover, most of the surplus generated by the CSI went to the wealthiest Zipcodes.

My work here has its limitations and can be improved by future research. For example, I did not consider the relationship between markets across time and location, and I assume that each Zipcode-month market is independent. I plan to develop a structural model that would consider a consumer's decision on whether and when to purchase solar panels. Nearby markets may also be correlated, since some sellers may operate in several adjacent Zipcodes.

Another model with one unified supply market but disaggregated demand market can be used to test this theory. I can't yet identify whether the impact of the CSI program is long-run or merely short-run. It could be that the rebate just incentivized households to adopt a solar-panel system earlier than they intended to. Despite this paper's limitations, my analysis here of the California solar market could help guide policy makers in designing future incentive programs.

CHAPTER 2

Estimates of the Impact of Pasadena's Minimum Wage Ordinance

Edward E. Leamer, Jonathan Gu, Mengshan Cui

2.1 Introduction

California Senate Bill No. 3, which was approved by Governor Brown on April 4, 2016, established a California minimum wage equal to $10.50 per hour for employers with 26 or more employees beginning on January 1, 2017, and stipulated annual increases in the California minimum wage up to $15 per hour on January 1, 2022. Prior to the passage of the California minimum wage, the City of Los Angeles had legislated its own minimum wage schedule with a level of $10.50 on July 1, 2016, six months earlier than the State of California, with increments that increase the LA City minimum wage to $15 on July 1, 2020, a year and a half before the California State minimum wage will reach $15.

The Pasadena Minimum Wage ordinance (Ordinance 7278) passed on March 14, 2016 adopts the City of LA minimum wage schedule through the end of June 2019. This paper studies the impact of minimum wage on important economic variables. Furthermore, we would like to provide policy implications of what are the possible benefits and risks of continuing on LA city's higher minimum wage schedule versus going back to California state minimum wage schedule. We have worked hard to distinguish the effect of the California minimum wage increases from the Pasadena increment since the City of Pasadena cannot call off the future increases in the California minimum wage and thus has discretion over only it's local increment. This is not easy to accomplish because the evidence so far is quite limited.

Figure 2.1: Pasadena Colorado Boulevard

An example of something that might be at stake in this local minimum wage decision is the location of restaurants along Colorado Boulevard illustrated in Figure 2.1. Colorado St/Blvd extends from Glendale through Eagle Rock and into Pasadena, with restaurants on all three segments. The Eagle Rock segment is governed by the higher minimum wage of the City of Los Angeles but Glendale has the lower minimum wage of the State of California. Eagle Rock may have the most at stake here, since if the City of Pasadena opts for the lower minimum wages of the State of California then Eagle Rock would face lower-wage competition both from the East (Pasadena) and from the West (Glendale), and jobs customers could move from Eagle Rock into both Pasadena and Glendale. On the other hand, if Pasadena continues to opt for the high-minimum-wage schedule of the City of Los Angeles, that puts enterprises within Pasadena in an adverse position compared with places like Glendale, La Cańada Flintridge and Alhambra and Monterey Park. The very limited experience with the Pasadena increment so far has not produced evidence of this kind of movement of jobs or enterprises, but the difference between the City and the State minimum wages is going to be larger in the years ahead, with possibly more impact.

The work described in this document is based primarily on the Quarterly Census of Employment and Wages collected by the State of California. We use these data to assess the impact of the California and Pasadena minimum wages on number of establishments, number of employees, and earnings per employee. We also use Pasadena and Los Angeles sales tax revenue to carry out a similar analysis to determine the impact of the minimum wage on sales tax revenue.

Solid conclusions regarding the impact of Pasadena minimum wages from 2011 to 2018 on earnings, employment, and number of establishments are difficult to make because of the limitations of the minimum wage "experiment" that has so far occurred, because the data we rely on only has labor earnings and number of workers but not hours worked, because the data are not individuals but enterprise based, because the geography of temporarily lower minimum wages surrounding Pasadena is complex, because the California minimum wage legislation dictates the precise dates when some workers must receive their wage increases but all other responses to this legislation may be made slowly over time possibly in anticipation of higher minimum wages to come, and because each industry has unobserved drivers that might mask the effects of minimum wage increments.

However, using several different econometric models for interpreting the data from 2011 to 2018, the evidence overall points to a positive impact of California/Pasadena minimum wages on the earnings of restaurant workers and of other low wage industries, confirming that the law is being obeyed. Our preferred model implies that a minimum wage increase of 10% would increase the average quarterly earnings per worker in limited-service restaurants by 8% and in full-service restaurants by 5%[1].

Our model also supports the conclusion that about half of the total increase in earnings resulting from a minimum wage increase occurs within the first quarter of the minimum wage increment. This response is consistent with the legislation which directly and immediately

[1]This increase in average earnings does not mean necessarily that the low-wage workers are better off. An increase in earnings per worker might occur if the workers with the lowest earnings were laid off but we have not found evidence of job losses coincident with the earnings increases. It is also possible that the increase in average earnings per employee is a result of a reduction of hours worked by the low-wage employees and/or an increase in hours worked by the high-wage employees. Absent data on hours worked we are not in a position to comment on this possibility.

affects only part of each firm's employees but has lingering effects on the others.

While effects on average wages of employed workers are clear in the theory and clear in the data, employment effects are not a sure thing theoretically and are harder to detect in the data. The economists' favorite supply and demand model makes it a virtual certainty that job losses come with minimum wage increases. It is only a matter of when and how much. But there are two other theoretical reasons why employment effects could be absent. One theory is that wages are determined not by competitive labor markets but by bilateral bargains between employers with many options and employees with few. For that reason a minimum wage might improve the bargaining power of workers and support higher wages with no loss of employment.The second theoretical reason why there may be small employment effects is that industry-wide increases in costs are normally passed on to customers in the form of higher prices. If these higher prices do not reduce sales, the level of employment required to provide those services also remains the same.

This discussion of the theory of employment effects of the minimum wage foreshadows the fact the evidence about employment effects is not so clear. Our preferred model only shows convincing negative employment effects of a minimum wage increase local to Pasadena for Limited Service Restaurants. Overall the traditional error bands around our estimates of the impact of either the State minimum wage and the Pasadena minimum wage on the 24 industries within our dataset are wide enough to include zero. To express this differently, the employment response to higher minimum wages is neither so sudden nor so great to make it transparent in the data we are studying, though a negative employment response appears present when viewed with the help of some models.

This work is closely related with literature that study the impact of minimum wage in Los Angeles. Both Beacon Economics (Thornberg et al. (2015)) and Reich et al. (2014) from UC Berkeley Institute for Research on Labor and Employment constructed reports for the City of Los Angeles evaluating the impact of future minimum wage increase on the workers, business, and economics in the city. However, these two reports make conflicting conclusions.

The Berkeley group uses county level data to derive that increasing minimum wages to be beneficial to the city by increasing worker's earning. This study also finds no significant employment effect due to the minimum wage increase. The report by Beacon Economics argues that minimum wage increases would fail on a cost-benefit basis. This report uses ACS data and estimation result from Meer and West (2016) to conclude that only one in every four dollars of increasing cost goes to low income workers.

Literature that studies minimum wage have conflicting results in general. Meer and West (2016) finds that minimum wages reduces job growth. Jardim et al. (2018) evaluates the effects of the Seattle minimum wage ordinance. It concludes that a higher minimum wage reduced hours worked in low-wage jobs by 9%, while hourly wages in such jobs increased by 3% . Baskaya and Rubinstein (2012) and Congressional budget Office (CBO (2014)) both studied the impact of federal minimum wage. The former one finds substantial disemployment effect of minimum wages on teenagers, while the latter one finds increase in earnings would not go to low income families. Sabia et al. (2012) estimate the effect of the New York State minimum wage increase. They found the minimum wage increases is associated with 20.2% to 21.8% decrease in the employment of low-skilled workers.

Other literatures such as Dude et al. (2010), Allegretto et al. (2011), Allegretto et al. (2017), Card and Krueger (1994), Dude et al. (2007), Addison et al. (2009), Giuliano (2013), and Hirsch et al. (2015) find no effect of minimum wage on employment. Neumark and Wascher (2011) study the effects of the interactions between the Earned Income Tax Credit and minimum wages on labor market outcomes. Wicks-Lim (2006) documents ripple effects of minimum wage. Brochu and Green (2013) uses Canadian data to study the labor market transition rate.

There are relatively few research papers that examine a minimum increase at a scale so specific to one city. Our research uses data from the individual zipcodes within and around Pasadena to examine the impact of minimum wage. The study that closest matches our detailed geographic examination is the study by Jardim et. al (2017), but this study still uses data that compares the employment across two counties. However, there could be large variance across cities within a county and a wealthier city with higher average wages and home prices may be less susceptible to disemployment effects. There are areas in Pasadena that are filled with wealthy households, and there are also zipcodes within Pasadena that are filled with college students attending Caltech or other institutions. We break the city of Pasadena down into five different zones based on the average level of income. For each zone we create a comparison group from the zipcodes in the surrounding region that most closely mimics the conditions of the region around Pasadena. The impact of minimum wage could be different between these areas. However our data does not offer any significant differences in impact by the different wealth levels within Pasadena.

Furthermore, the impact of minimum wages can be broken down by industry as well. Many recent studies have focused on the fast-food industry, and indeed we do as well, but we also highlight the potential impact of minimum wages on other vulnerable industries as well. For example, our data shows that the opening of new hair, nail, and skin care services have dropped off in Pasadena since the minimum wage increase.

The rest of the paper is organized as follows: Section 2.2 describes the minimum wage schedule in California, Los Angeles County, and Pasadena. We further discuss the limitation of this study cause by the policy design. Section 2.3 contains data source and summary statistics. Section 2.4 discusses main model and specification. Section 2.5 presents findings from the main model. Section 2.6 offers robustness check. Section 2.7 concludes.

2.2 Minimum Wage Schedule

An ideal minimum wage experiment would be a randomized controlled trial in which a group of identical regions is randomly divided into two groups: one group with an increase in the minimum wage and the other with no increase. Then the data on employment, for example, can be summarized in four numbers: the levels of employment in the two groups, both before and after the minimum wage increase. If the communities that experienced the minimum wage increase had a smaller increment in employment than the communities that did not have the minimum wage increase, we would conclude that the minimum wage was suppressing employment. That is what economist call a “difference in differences” estimate.

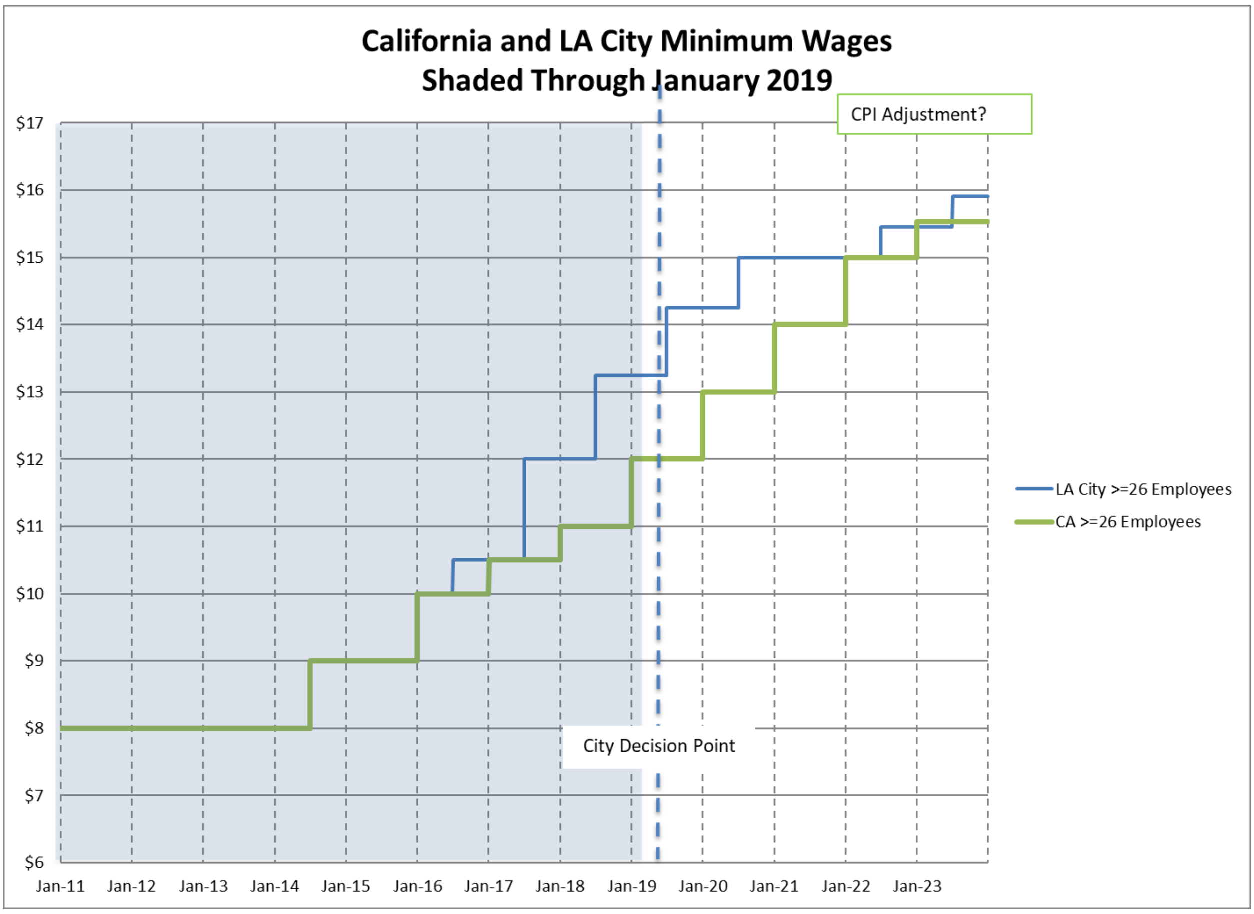
Unfortunately, there are no such experiments. There are no identical regions that have adopted different minimum wages. The level of local minimum wage was never chosen randomly but was determined by a political process that is presumably sensitive to the possibility that a minimum wage set too high can have adverse employment outcomes. If we discover that the sickest people take the most medicine, that is not proving that the medicine has adverse effects. Likewise, if we discover that the communities with the highest minimum wages have the greatest increases in employment, that is not proving that higher minimum wages increase employment. What we are saying is that it's complicated to pull from the data convincing evidence about the effects of the minimum wage. But we have to do the best with what we have, providing appropriate caveats when needed. The first step in that journey is to think clearly about the nature of the experiment we are observing.

We think that the two major problems with the data that we have available are: (1) the whole schedule of minimum wage increases was announced in advance, allowing firms to react in anticipation of minimum wage increments yet to come. (2) the Pasadena minimum wage increment creates a complex local geography of business competition, allowing enterprises to escape the Pasadena increment with a fairly short move to a different jurisdiction. These two issues are now discussed.

2.2.1 The Minimum Wage Increases are determined years in advance

The California and City of LA minimum wage schedules beginning in 2011 (the first year of the Pasadena data that we are studying) are illustrated in the Figure 2.2 which has a shaded region representing the history ending in 2019Q1, the last quarter of our data, and a dashed vertical line indicating the limit of Ordinance 7278, at which point Pasadena will either revert to the California minimum or stick with the LA minimum or something else.

Figure 2.2: California and City of Los Angeles Minimum Wages



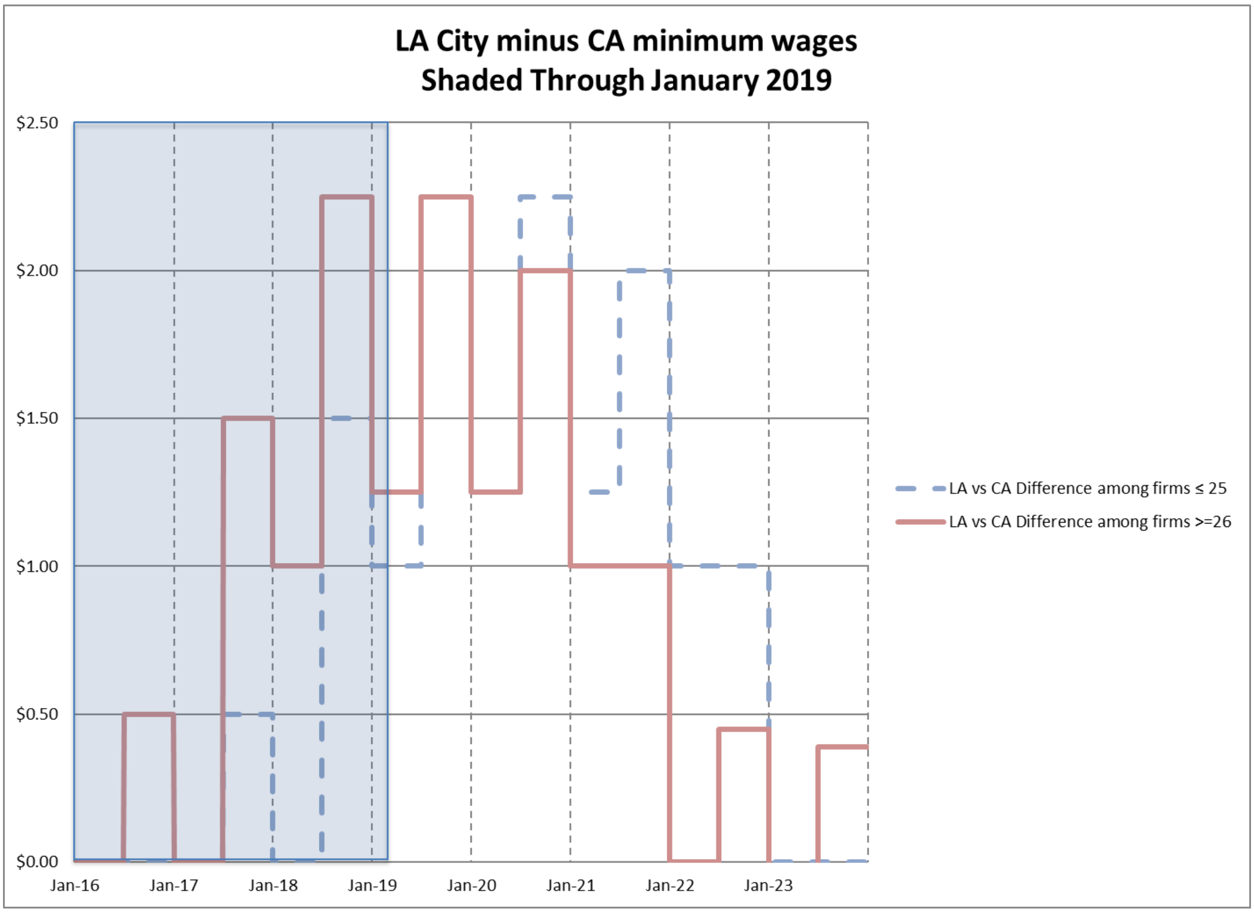
The legislation adopted by the State of California and by the City of Los Angeles firmly established increases in the minimum wage for six or seven years into the future and even indefinitely because of the inflation adjustment that commences in 2022/2023. The best way to summarize this graph in words is that California and Los Angeles/Pasadena have adopted two different but parallel paths toward $15, which means that the impact of the Pasadena ordinance might be only to accelerate by a year or two the impact of the California minimum wage.

But it's more complicated than that. This legislation gives businesses plenty of advance warning and plenty of time to plan how to respond, such as by moving to another location or not opening a new enterprise, by changing the nature of the service provided, by adopting human resources systems that weed out the less productive workers, by automating, by passing the incremental costs on to customers via higher prices or onto building owners via lower rents, or by owners absorbing part of the cost increase. The possible reactions are quite diverse and many are hard or impossible to identify in the data that we have. In particular, it may be difficult to identify an employment effect because any employment reductions that occur can be more a consequence of the whole schedule of minimum wage increments rather than the year-by-year increments. The data analysis that we carry out focuses on the year-by-year increments and only incidentally picks up the effect of the whole schedule. This is quite different from the likely evidence about wage effects since the legislation stipulates exactly when wages have to increment, which is something we should be able to see in the data, and do.

2.2.2 The Local Increment to the California Minimum Wage is Small and Variable

We will be studying the possibility that the Pasadena increment has a different effect than the California minimum wage. Our models will include two variables: (1) the prevailing minimum wage equal to the California minimum wage plus the local increment and (2) the local increment which is the amount by which the Pasadena minimum wage exceeds the California minimum wage. The second variable has a zero coefficient if all that matters is the prevailing minimum wage but a nonzero coefficient if the effect of the local increment is different. For wages we expect the first coefficient to be positive and the second zero, meaning that what matters for setting wages is the prevailing minimum wage not how much of it is dictated by local legislation. For employment, we expect negative coefficients on both, meaning the adverse employment effect is greater for the local components of the minimum wage because it encourages firms to move to close locations with lower minimum wages. In contrast, escaping the California minimum wage requires a move out-of-state. On the other hand, moving from Pasadena to one of the surrounding communities would only delay the minimum wage increment by about a year and a half, and that short delay might not justify the cost of moving. In that case, responses like automation at the Pasadena location might be preferred to moving in pursuit of a temporarily lower minimum wage.

Figure 2.3: City of Los Angeles Increments to the California Minimum Wage



The local increment for the City of Los Angeles is illustrated in Figure 2.3, which distinguishes enterprises with more than 25 employees from smaller enterprises. Here we see a problem for our study: through January 2019 the Pasadena increment was only $0.50 in the second half of 2016 and $1.50 for the second half of 2017 and then $2.25 in the second half of 2018 for firms with 26 or more employees, but much less for firms with 25 or fewer employees. That difference should show up in wages but maybe not so clearly in employment.

2.2.3 Geographical Variability of Minimum Wages

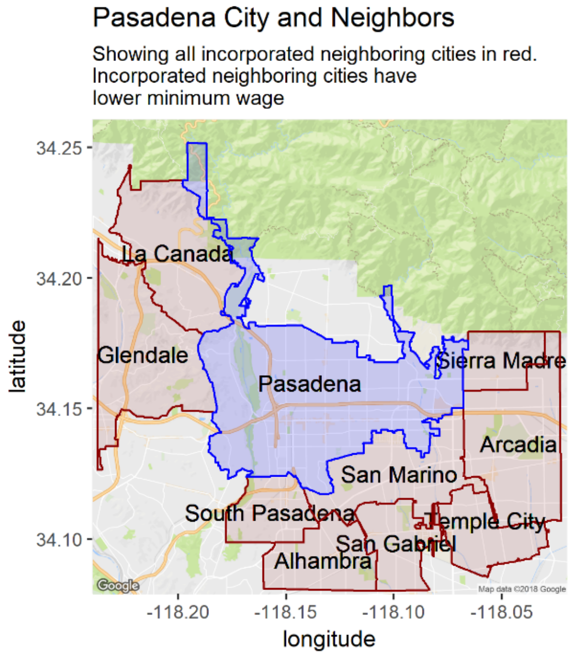
The Pasadena/City of LA increment to minimum wages creates a geographical aspect to the minimum wage experiment by establishing adjacent or close communities with different minimum wages. The local geography is illustrated in the four images in Figure 2.4. Figure 2.4(a) has Pasadena shaded in blue and adjacent or close regions that are subject only to the California minimum wage shaded in light red. (La Canada, Glendale, South Pasadena, Alhambra, San Gabriel, Temple City, San Marino, Arcadia and Sierra Madre.) The lighter regions to the northeast and southwest of Pasadena are Altadena and the City of LA, both with the same minimum wage schedule as Pasadena.

A special risk created by the Pasadena minimum wage is that jobs might leave Pasadena in favor of one of the close cities with a lower minimum wage. That could make the effect of the Pasadena increment on employment greater than the effects of the California increments. It also raises the possibility that we will double-count the employment effects if we use regions close to Pasadena as a control group for Pasadena since we would count the job loss in Pasadena and also the job gain in the neighboring community.

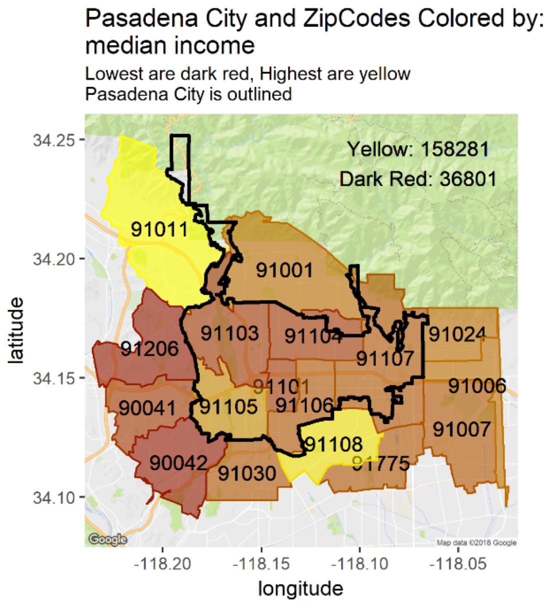
This image captures the difficult question that confronts the Pasadena City Council: should Pasadena align itself with the City of LA and Altadena, which would encourage the movement of jobs to the region shaded red (Glendale, La Canada, South Pasadena, San Marino and so on), or should Pasadena align itself with the red region, thus encouraging a job flow into Pasadena or other red cities out of the City of LA and Altadena.

The three other images in Figure 2.4 help understand what is at risk in this decision. Figure 2.4(b) has the zip codes color-coded by median income of the residents. The highest median incomes are in La Cańada Flintridge and San Marino. Within Pasadena the south-western zipcode 91105 has a high median income but the rest of the zipcodes have lower and comparable income levels. Figure 2.4(c) illustrates the percent of the residents who work in food service and accommodations. It is the northern zipcodes of Pasadena, 91103 and 91104, that have high fractions of residents in this sector. Outside of Pasadena the region

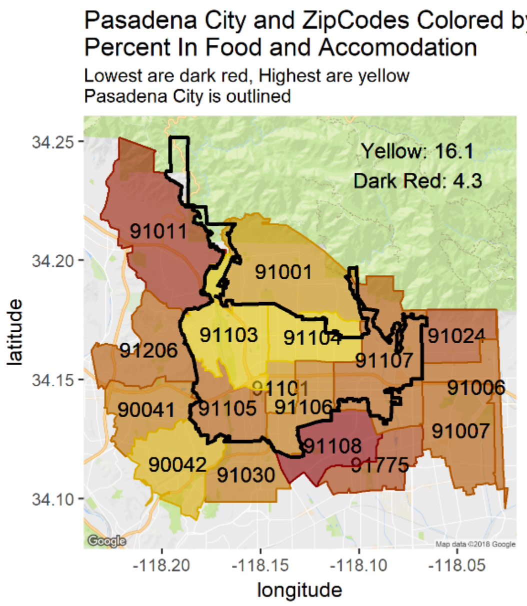
Figure 2.4: Map of Pasadena and surrounding cities



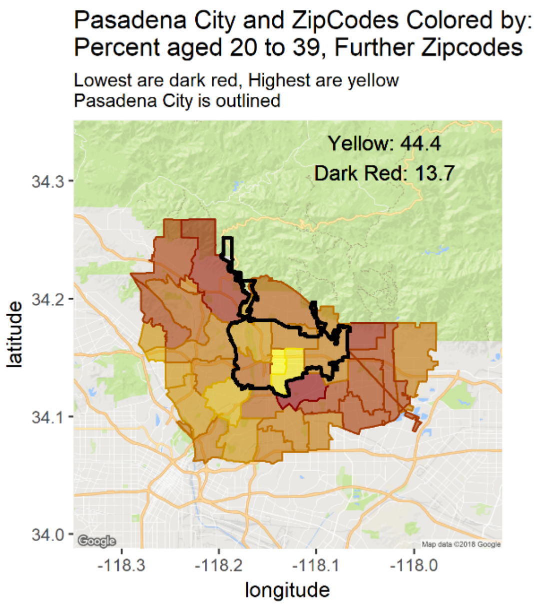
(a)



(b)



(c)



(d)

with a high fraction of the residents in food services and accommodations is Highland Park (90042).

Another geographic complexity is that Pasadena has neighborhoods that are quite different in terms of income, age, and sectoral job mix. Per the data reported in the Table 2.1, median incomes within Pasadena vary from a low of $61,473 in 91101 to a high of $107,284 in 91105. Among the other differences are: 48.7% of workers in 91101 were young (20-39) while 27% were young in 91105; 27% earned less than $25,000 in 91101 but only 11.1% in 91105. It is likely that the younger lower-paid workers from 91101 would be more impacted by the minimum wage than older better paid workers who live in 91105, but our data sets are based on location of work not location of residence.

Table 2.1: Check for Balanced Characteristics of Comparison Groups

Group		1			2	
city	Glendale	Alhambra	Pasadena*	Temple City	Monrovia	Pasadena*
Zipcode	91202	91803	91101	91780	91016	91103
Total Population	23219	29502	20761	35674	41901	28124
Number of Households	8768	9566	10745	11305	14699	8381
Median Income	62104	57380	61473	62461	67868	62697
Age 20-39	33.8%	32.3%	48.7%	29.2%	32.7%	36.4%
High School or less	23.2%	32.7%	13.8%	26.9%	26.5%	28%
Earning less than $25,000	23.6%	21.5%	27.4%	20.2%	17.1%	23.4%
Labor Force Participation	61%	60.7%	68.3%	59.1%	71.1%	63.6%
Unemployment rate	8.7%	5%	7.2%	7%	9.3%	7.2%
Occ—Ind						
Service	14.2%	21.9%	11.9%	16.9%	18.1%	25.1%
Sales	29.1%	26.9%	17.9%	32.3%	24.7%	21.4%
Construction	3.2%	4.8%	4.3%	4.4%	6.1%	8.6%
Retail	12.7%	11.2%	5.9%	11%	10.5%	10.3%
Accommodation and Food	6.9%	12.3%	10.4%	10.9%	10.6%	13.2%
Group			3			
city	Arcadia	Montrose	Pasadena*	Pasadena*		
Zipcode	91007	91020	91106	91104		
Total Population	34619	8448	24875	38725		

Continued on next page

Table 2.1 – continued from previous page

Number of Households	11647	3345	10540	13081		
Median Income	75353	70014	75160	70208		
Age 20-39	25.3%	33.6%	44.9%	33%		
High School or less	21.7%	20.8%	12.5%	22.3%		
Earning less than $25,000	17.7%	18.7%	16.9%	21.6%		
Labor Force Participation	58.5%	68%	70.4%	66.2%		
Unemployment rate	7%	7.4%	5.5%	8.5%		
Occ—Ind						
Service	11.9%	11.8%	11.7%	19.7%		
Sales	28%	28.5%	18.4%	20%		
Construction	3%	3.8%	3.8%	3.9%		
Retail	8.6%	8.5%	8.1%	8.6%		
Accommodation and Food	8.6%	6.3%	10%	13.3%		
Group		4			5	
city	South Pasadena	San Gabriel	Pasadena*	Glendale	Sierra Madre	Pasadena*
Zipcode	91030	91775	91107	91208	91024	91105
Total Population	25905	25389	32027	17180	11067	11728
Number of Households	10150	8164	12502	5876	4403	5485
Median Income	84683	79637	84663	111563	95256	107284
Age 20-39	30.7%	27.2%	31.3%	28.7%	23.1%	27%
High School or less	11.8%	24.6%	15%	14.4%	11.6%	10.8%
Earning less than $25,000	13.2%	15.6%	15.1%	9.6%	10.7%	11.1%
Labor Force Participation	70.8%	61.2%	64.9%	65.3%	66.2%	64.5%
Unemployment rate	6.1%	3.8%	6.6%	4.5%	5.3%	5.9%
Occ—Ind						
Service	10%	13.8%	12.3%	11.3%	5.5%	6.9%
Sales	19.9%	23.1%	23%	25.9%	25.1%	17.8%
Construction	4%	4.6%	4.2%	3.9%	3.4%	5.3%
Retail	6.7%	8.9%	9.5%	8.5%	7.5%	5.6%
Accommodation and Food	9.3%	6.9%	8.3%	9.1%	6.7%	8.7%

*: Above State Minimum Wage

2.3 Data

2.3.1 Data Source

Two main data sets we use are Quarterly Census of Employment and Wages and Sales Tax Revenue.

We rely primarily on data collected by the Quarterly Census of Employment and Wages. Every enterprise in the United States is required to report quarterly the total wages paid in the quarter and the number of employees in each month of the quarter.

The sales tax data has been assembled by HdL Companies and contains quarterly city level data for sales tax revenue for apparel, fast casual dining, casual dining, quick-service dining, and specialty stores. This data set includes the city of Pasadena, Glendale, Monrovia, Burbank, Arcadia, Temple City, Sierra Madre, West Hollywood, Santa Monica, the city of Los Angeles, and Los Angeles County. It covers the period from 2011 quarter 1 to 2018 quarter 1.

Other data sets includes American Community Survey and Current Population Survey.

2.3.2 Group

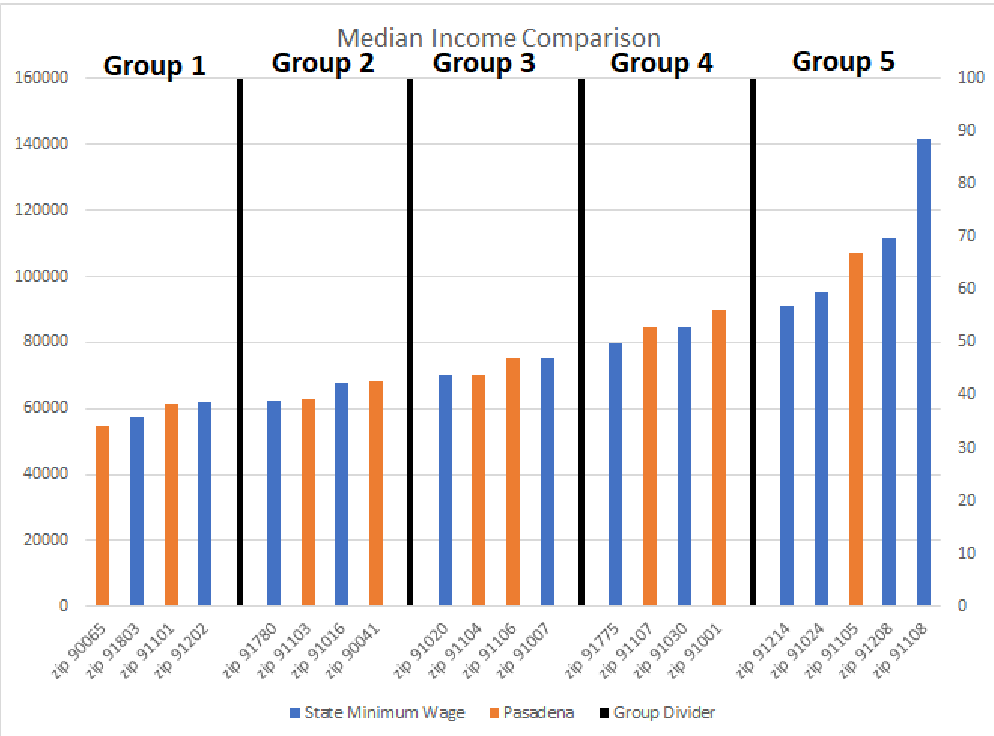
Our strategy for estimating the impact of the Pasadena minimum wage is to compare pairs of regions that are similar to each other but have different minimum wage schedules. We have split Pasadena and its surrounding regions up into ten areas, with five areas consisting of a distinct section of Pasadena and five areas capturing economically similar areas around Pasadena. These areas capture much variation in income within the Pasadena: for example the neighborhood of Pasadena to the southeast near San Marino is quite wealthy, and we would like to compare this wealthy Pasadena neighborhood with another relatively wealthy district nearby that is not impacted by the Pasadena minimum wage ordinance. As another example, the area around Cal tech is populated by many residents between the ages of twenty and thirty, and we would have found two other zipcodes near Pasadena that has the most similar economic and demographic characteristics. Several of the groups also include close zipcodes outside Pasadena with the same minimum wage as the Pasadena zipcodes.

Table 2.2: Zipcodes with similar median incomes

Group	Far Option
G1:	Low MW: Alhambra 91803, Glendale 91202 High MW: Pasadena 91101, LA 90065
G2:	Low MW: Temple City 91780, Monrovia 91016 High MW: Pasadena 91103, LA 90041
G3:	Low MW: Montrose 91020 Arcadia 91007 High MW: Pasadena 91104, 91106
G4:	Low MW: San Gabriel 91775, South Pasadena 91030 High MW: Pasadena 91107, Altadena 91001
G5:	Low MW: Sierra Madre 91024, Glendale 91208, San Marino 91108, La Crescenta 91214 High MW: Pasadena 91105

Our groups are reported in Table 2.2 which begins with Group 1 which has a high minimum wage region composed of Pasadena 91101 and the City of LA 90065, contrasted with the low MW zipcodes in Alhambra and Glendale. Figure 2.5 illustrates the median incomes in each of these zipcodes by groups, which was the basis for our groups. We try to

Figure 2.5: Median Income Comparisons



group zipcodes with similar income together. Table 2.1 provides balanced checks of control zipcodes (CA minimum wage) and treatment zipcodes(Pasadena and LA city). For the balance checks, we examine variables that are relevant to the impact of minimum wage. In Group 1, Pasadena 91101 (which surrounds Caltech) has more young people, more educated people, more people earning less than 25,000. In terms of occupation, Pasadena has less Service, Sales, and Retail than their proposed controls in Alhambra and Glendale. In Group 2, Pasadena 91003 has more young people, but also it has more less educated people and people with low earnings. Here we can see the benefits of including more zipcodes. Temple City and Monrovia are large zipcodes with population above the median of our sample. In terms of occupation, Pasadena has more service and less sales. In Group 3, we have a very small proposed control zipcode in Montrose. Montrose is tiny city, with only one zipcode and a population of 8500. Pasadena 91106 has more young people and is more educated than the proposed controls. Pasadena 91104 is actually quite similar to the proposed controls, although it twice the percentage of people working in service occupations. Among the rich counties, all the zipcodes are quite similar in terms of characteristics are a likely to affect the impact of minimum wage.

Figure 2.6: Map of Five Comparison Groups

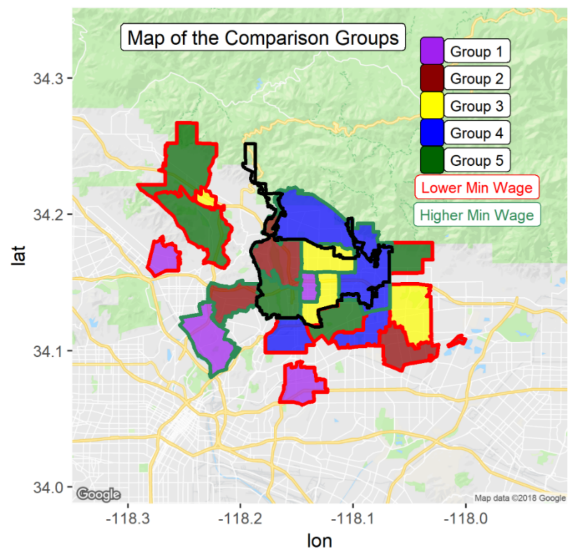


Figure 2.6 is a color coded map of these regions. If we were to do a complete local city comparison, we would simply compare the blue regions with the red regions. Further analysis show that there is strong zipcode level heterogeneity within the cities. We would be better off comparing zipcodes that are similar with each other. Figure 2.4(b) shows the variation in income. We can see the

pitfalls of comparing the zipcode 91105 in Pasadena with zipcode 91206 in Glendale. The Pasadena zipcode has a much higher median income. Figure 2.4(c) shows that Pasadena zipcodes 91103 and 91104 have the highest percentages of people working in the food and accommodation occupation. Finally we can see that Pasadena zipcodes 91101 and 91106 have 44% of their population aged between 20 and 40. For reference, classic Old Town Pasadena and Caltech are in zipcode 91101. The administrative buildings and dormitories of Caltech actually have their own zipcode (91126).

Figure 2.7 presents city (zipcodes within a group with same minimum wage schedule) industry composition difference. We present the top-ten employed industries from QCEW non-confidential data. The "other" means all other industries that are non-confidential. This

Figure 2.7: Group Industry Composition

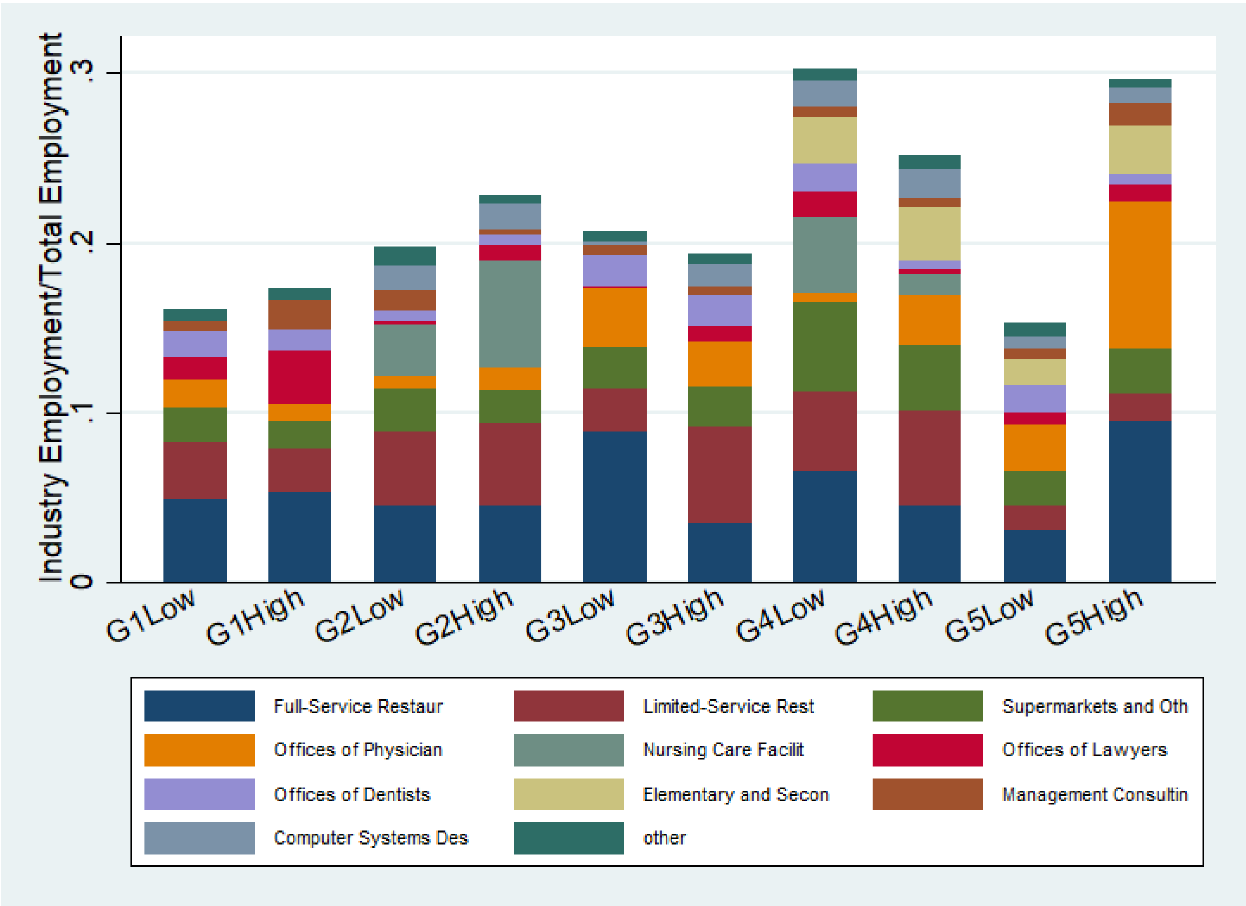


figure shows that cities within the same group have relatively similar industry compositions. However, each city has very different industry compositions. For example, Group 1 cities have relatively higher ratio of employments in offices of lawyers without any employment in nursing care or computer system. Group 2 cities have high level of employment in nursing care facilities. Group 4 and 5 have higher employment in Elementary and Secondary Schools. This figure emphasizes the importance of including a city-industry fixed effect in regression to better control for industry composition difference.

2.3.3 Summary Statistics

Table 2.3 includes all the industries for which the Quarterly Census of Employment and Wages has data for going back to 2011. The values are average employment numbers during the period 2011 quarter 1 to 2017 quarter 4 for the region with the Pasadena minimum wage composed of Pasadena, Altadena zip code 91001, and LA zip codes 90041, 90065. The sectors are sorted by employment levels and each column shaded with the largest numbers dark and the smallest light.

Figure 2.8 illustrates the fractions of minimum wage workers in various industries. At the top are hair and nail salons with 60% of the workers paid less than $12 per hour, and restaurants with 50% of their workers in that category. These are sectors which require special scrutiny.

Table 2.4 compares the sales tax revenue in the whole of Los Angeles county with the City of Pasadena in year 2011 and 2017. From this table we can see that the biggest source of sales tax revenue in the county is quick-service dining with around $66 million in sales taxes in 2011 and almost $93 million in 2017. However, in Pasadena, both casual dining and apparel have larger sales than quick-service dining in both 2011 and 2017. However, quick-service dining grew 41.49% in Pasadena from 2011 to 2017, while apparel has almost no growth during this period. The standout industry in terms of growth of revenue in both LA County and Pasadena is fast casual dining. From this table, it does not appear that the

Table 2.3: Pasadena Industry Detail

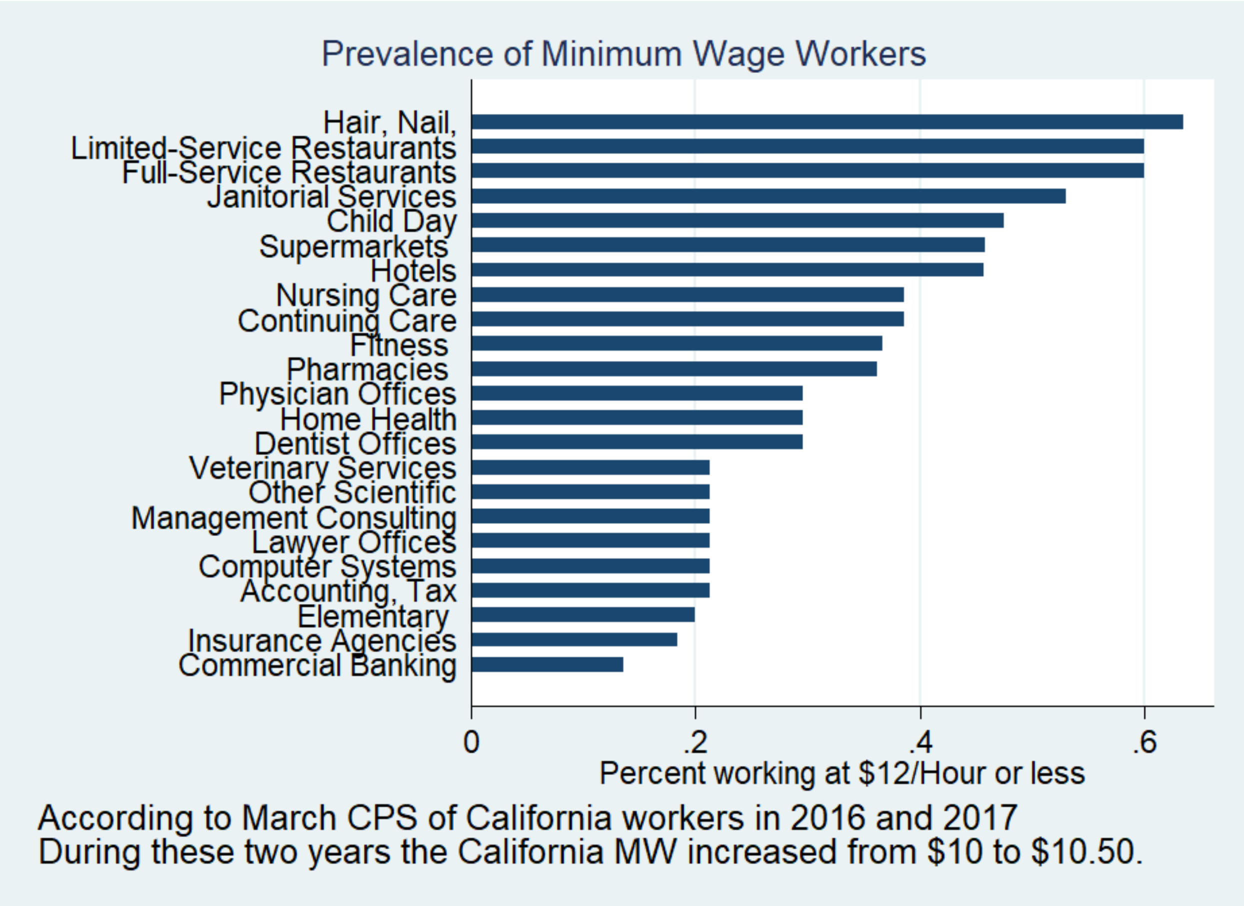
Industry	Employment	Firms	Earnings Per Person Per Quarter
Full-Service Restaurants	6361	257	\$6,379
Limited-Service Restaurants	4662	235	\$5,298
Physician Offices	3139	502	\$17,531
Supermarkets and Groceries	2488	42	\$7,385
Lawyer Offices	1550	401	\$20,082
Elementary and Secondary Schools	1441	24	\$13,627
Nursing Facilities	1363	19	\$8,364
Computer Systems Design	1215	130	\$23,756
Management Consulting Services	1126	187	\$22,024
Dentist Offices	1060	173	\$11,545
Insurance Agencies and Brokerages	1017	138	\$20,631
Accounting, Tax Preparation	759	109	\$14,411
Pharmacies and Drug Stores	583	61	\$11,442
Child Day Care Services	563	49	\$6,858
Residential Building Construction	560	114	\$14,925
Hair, Nail, and Skin Care Services	539	105	\$5,240
Home Health Care Services	475	16	\$9,179
Other Technical Consulting Services	264	138	\$16,373
Veterinary Services	256	22	\$8,979
Commercial Banking	237	21	\$17,635
Fitness and Recreational Sports Centers	212	10	\$4,769
Hotels and Motels	188	16	\$5,416
Continuing Care Retirement Communities	185	5	\$6,073
Janitorial Services	61	11	\$5,894

increases in minimum wages are reducing tax revenue, but more on this below when this data is filtered through an econometric model.

Table 2.4: Industry Annual Sales Tax Revenue Comparasion between LA county and Pasadena

	2011		2017		Growth Rate 2011-2017	
Industry	LA County	Pasadena	LA County	Pasadena	LA County	Pasadena
Quick-Service Dining	\$66,455,360	\$1,084,072	\$92,918,480	\$1,533,899	39.82%	41.49%
Apparel	\$66,382,680	\$1,755,670	\$84,285,120	\$1,767,651	26.97%	0.68%
Casual Dining	\$54,843,800	\$2,027,759	\$86,950,080	\$2,942,134	58.54%	45.09%
Specialty Stores	\$32,802,512	\$753,472	\$40,547,560	\$863,900	23.61%	14.66%
Fast Casual Dining	\$7,598,740	\$285,851	\$17,678,712	\$712,750	132.65%	149.34%

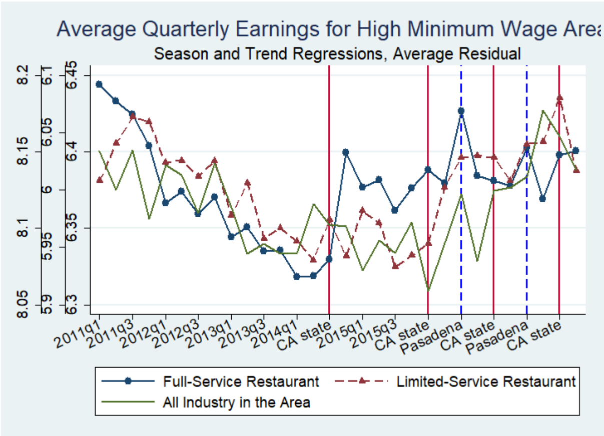
Figure 2.8: Prevalence of Minimum Wage Workers



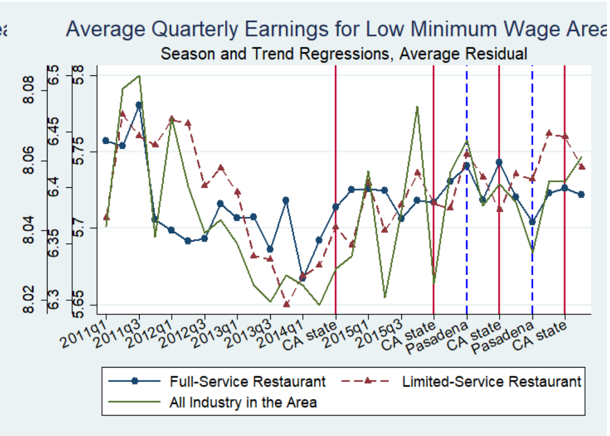
2.3.4 Industry Summary Statistics: Restaurants

In order to study the impact of minimum wage, we will focus on industries that are low skill labor intensive. Restaurant is one of the most important industry with low income workers. We further divide restaurants into limited and full service restaurants. Figure 2.9 present inter-temporal patterns (controlling for seasonal fixed effects and a time trend) of (1) average earnings, (2) employment, and (3) number of establishments. Each figure includes the data for all-industries, and for full-service and limited-service restaurants. Figures are presented for high minimum wage areas (Pasadena, and zipcodes 91001, 90041, and 90065 in Los Angeles, and Altadena) and for low minimum areas. All figures include vertical lines that indicate when either the California or the Pasadena minimum wage was increased.

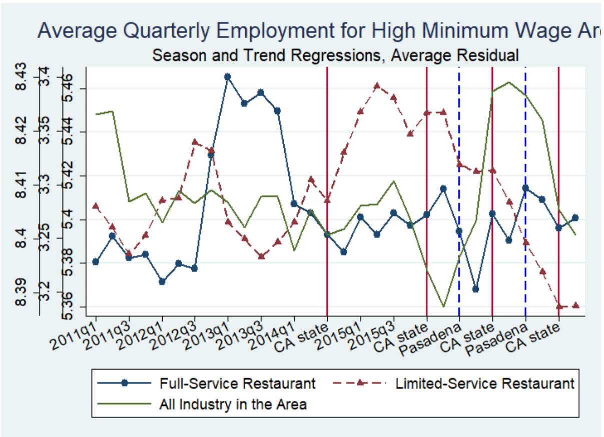
Figure 2.9: Average Quarterly Earnings, Employment, and Establishment for Restaurants



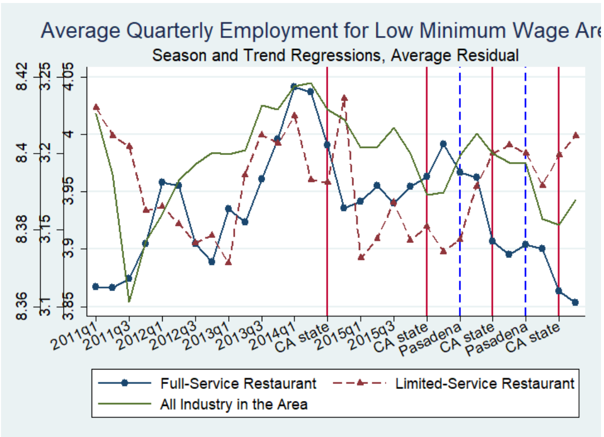
(a)



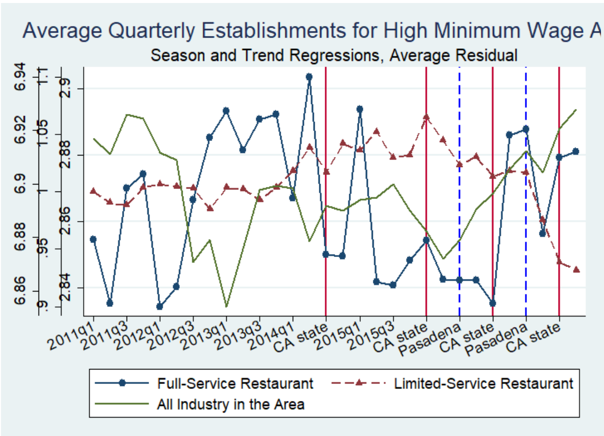
(b)



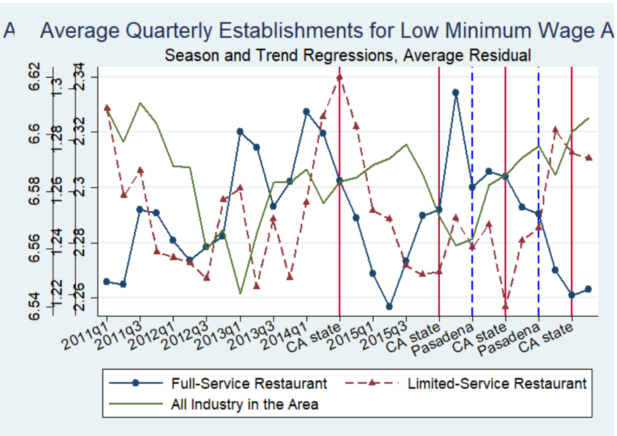
(c)



(d)



(e)



(f)

The removal of trends from all these figures supports visual displays that mimic the model-based analysis that also includes trends. These images are different if the trends are not removed, just as our estimates are different if the trend variables are not included. The main take-aways fro these figures are: (1) high minimum areas and low minimum areas have similar patterns for average of all industries, but very different patterns for restaurants. (2) Restaurants react to minimum wage changes very differently than the average of all industries. This indicates that minimum wages have heterogenous impact depending on the industry. The source of the heterogenous response in minimum wages could be due to the prevalence of minimum wage workers in each industry, whether low-wage workers are easy to substitute by technological capital, the average turnover of employees, etc. (3) Full and limited service restaurants react differently to minimum wage changes. This emphasizes the importance of looking at finer detail industry level. The finest detail that we have obtained from the QCEW is at the 5 digit NAICS level. The higher number of digits indicates a finer level of detail of the classification of businesses.

Figures 2.10 show seasonal adjusted quarterly fixed effect of

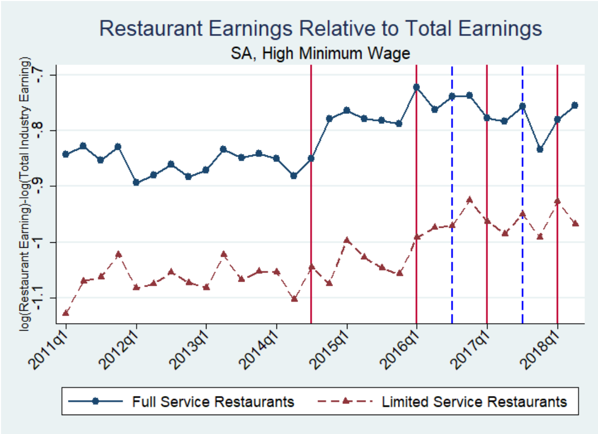
$$log(\frac{\text{restaurant earnings}}{\text{average earning of all industries in the area}})$$

and of

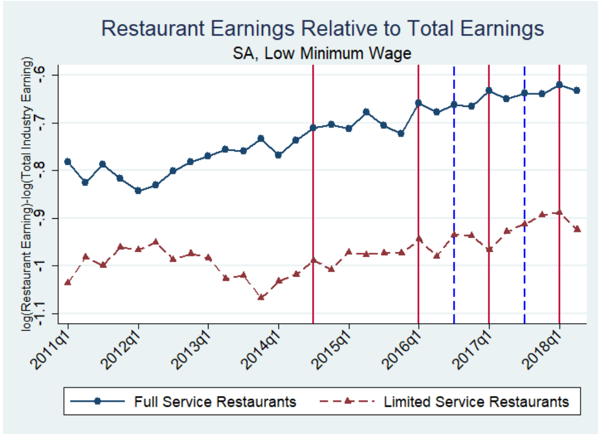
$$log(\frac{\text{restaurant employment}}{\text{employment of all industries in the area}}).$$

They are also presented separately for high and low minimum wage areas. A decreasing trend means the restaurant earnings or employment grows slower than the whole economy. An increasing trend means restaurant earnings or employment grows faster than the whole economy. The main messages from these figures are: (1) Restaurants earnings are slightly increased compared to the economy. The increase is more consistent after the first California minimum wage increase in July 2014. (2) Restaurants employment are increasing compared to the economy. This increase pattern seems to be unaffected by minimum wage change.

Figure 2.10: Restaurant Earnings, Employment, and Establishments Relative to Total

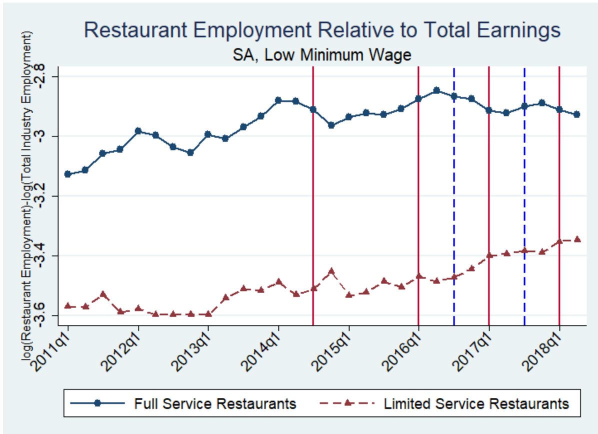


(a)



(b)

(c)



(d)

(3) Employment change pattern varies across high and low minimum wage areas and across industry. In high minimum areas, full-service restaurant employment moves very closely with the whole economy, while limited-service restaurant employment increases.

2.4 Models and Specification

The main regression being used in this research is:

$$\begin{aligned} log(y_{ict}) =& \beta_{0i} + \delta_i log(\text{City Minimum Wage}_{ct}) + \gamma_i log(\frac{\text{City Minimum Wage}}{\text{CA State Minimum Wage}}_{ct}) \\ &+ \beta_{1i} log(y_{ic(t-1)}) + \beta_{2i} log(Y_{ct}) + \beta_{3i} log(Y_{c(t-1)}) + \beta_{4i} Quarter_i + \theta_{si} + \epsilon_{ict} \end{aligned} \quad (2.1)$$

where:

- i: industry; c: city (zipcodes within same group with same minimum wage schedule); t: quarter.
- y_{ict} is the dependent variable for industry i in city c in quarter t.
- City Minimum Wage_{ct} is the minimum wage (per hour) that the city c is on in quarter t.
- CA State Minimum Wage_{ct} is the California minimum wage (per hour) in quarter t.
- $y_{ic(t-1)}$ is the dependent variable for industry i in city c in one quarter before.
- Y_{ct} is the total number of a variable in the city c in quarter t : $Y_{ct} = \sum_i Y_{ict}$.
- $Y_{c(t-1)} = \sum_i Y_{ic(t-1)}$.
- $Quarter_i$ represents industry level time trend.
- θ_{si} is being used to control for the industry-seasonal fixed effect.

The above regression is estimated for each industry separately. Our models use three "dependent" variables observed quarterly at the level of an industry in a particular region: Earnings per employee, Employment, and Number of Establishments. To explain the movements in these three dependent variables we have used "dynamic" models that allow the impact of an increment in the minimum wage to be spread over time. We include as explanatory variables two minimum wage variables, the prevailing minimum wage and the part of the prevailing wage that is due to the local legislation. We also include explanatory variables that reflect overall area-wide changes like the total employment and overall average earnings per employee which we take to be unaffected by minimum wages.

Each variable we have included in our models captures the effect one of the key factors mentioned above. Previous literature on minimum wage has mainly used the "two-way fixed effects" approach. Our model deviates from the previous literature in a number of ways, most notably, by taking into account the dynamic nature of our data: we are able to say how much of the impact of minimum wage we expect to occur in the first quarter. This difference is essential when analyzing dynamic data with measurements of the same quantity (such as employment in Supermarkets in Pasadena) over multiple periods. Without taking into account the dynamic nature of the data, some other researchers may assume that the number of employees on the payroll at Ralphs on Monday is completely independent of the number of employees on the payroll on the following Tuesday. In order to account for the correlation between outcomes we have included lagged dependent variables. These lagged dependent variables will also tell us how much of the minimum wage impact is expected to occur in the first quarter of a minimum wage increase. To study the long run impact of minimum wage, we use the fact the in the long run stable state, $y_{ict} = y_{ic(t-1)}$ and rewrite the regression equation into:

$$\begin{aligned} log(y_{ict}) =& \frac{\beta_{0i}}{1-\beta_{1i}} + \frac{\delta_i}{1-\beta_{1i}} log(\text{City Minimum Wage}_{ct}) \\ &+ \frac{\gamma_i}{1-\beta_{1i}} log(\left\{\frac{\text{City Minimum Wage}}{\text{CA State Minimum Wage}}\right\}_{ct}) + \frac{\beta_{2i}}{1-\beta_{1i}} log(Y_{ct}) \\ &+ \frac{\beta_{3i}}{1-\beta_{1i}} log(Y_{c(t-1)}) + \frac{\beta_{4i}}{1-\beta_{1i}} Quarter_i + \frac{\theta_{si}}{1-\beta_{1i}} + \epsilon_{ict} \end{aligned} \tag{2.2}$$

where $\dfrac{\delta_i}{1-\beta_{1i}}$ measures the long-run impact of minimum wage increase and $\dfrac{\gamma_i}{1-\beta_{1i}}$ estimates long-run impact of Pasadena increase its minimum wage above the state level.

In order to account for underlying forces that affect out outcomes separate from the minimum wage we have included a time trend and also the sum total of the outcome variable across all industries. The sum total outcome variable (such as the total number of employees in all industries in Pasadena) is included to reflect the changes in the economy that are local to the city.

The time trend is included to capture factors that may affect the real price of labor in the economy such as the constantly increasing technological progress, increasing availability of capital, or increasing rates of educated eligible workers. Without adding the time trend our results would actually be quite similar, indeed, without adding time trends we do find more results that are individually significant. However without a time trend the minimum wage is the only variable that documents the passage of time in our model, so any underlying force that is changing over the time of our study could be attributed to minimum wage, therefore we add time trends so that our results will indicate the impact of minimum wage above and beyond the time trend. As we can see in the data display of the number of establishments of hair, nail, and skin care services, including a time trend would lead us to expect that without minimum wages, the growth in the number of salons would have continued. This can be seen as both a positive and a negative attribute of the time trend: Positive if it were actually the case the hair, nail, and skin care services is a booming industry that would have continued it growth without minimum wages, and negative if we believe that the timing of number of salon establishments reaching an equilibrium level coincided with the implementation of California state minimum wage.

It is important to note that our analysis does give what we deem to be false positives because the industries that our model and our data report to be impacted by the minimum wage are not low-wage industries. Specifically, we see positive earnings impact of the minimum wage on veterinary services and dentist's offices even though the average employee at a veterinary clinic or a dentist's office makes twice as much as an average restaurant worker. These false positives highlight a caveat of our model: adding a linear time trend and total industry outcome variables into our model does not capture all of the underlying forces that can drive changes in earnings. If a sudden boom in dog ownership and dental hygiene occurred in 2014, then we cannot disentangle the sudden boom with the increasing California state minimum wage in 2014.

A third problem industry we have is the industry known as "Other Technical and Consulting Service" which is an amalgamation of consulting services that have not been classified into a specific industry. This sector is highly paid and ranks among the lowest in the proportion of employees that are working at minimum wage. This sector also happens to experience a nationwide decline in employment near the end of 2013, which precedes the California state minimum wage increase. This decline is likely simply a transfer of jobs from one industry code to another: on the aggregate level, there has actually been no change in the number of consulting jobs over this time, and management and business consulting (which have their own industry code) is on the rise during our data.

2.5 Main Results

2.5.1 Main Specification

For each industry and each dependent variable we have estimated a total of 24 different models. We report in this section the results generated by the one specification that we think yields the most reliable results. This model includes time trends, utilizes the data from all the five groups of regions together, and includes the Pasadena increment to the minimum wage.

We will first examine the findings of an increase in minimum wages inclusive of the Pasadena increment. It is important to note that these results may be driven primarily by

Table 2.5: Regression Result for Predicting Impact of Minimum Wage on Earnings per person

Industry	log(MW)	log(Incre*)	$y_{(t-1)}$	$log(Y_t)$	$log(Y_{t-1})$	Quarter	R^2
Accounting, Tax Preparation, Bookkeeping	0.306	0.698	0.242	0.265	0.583	-0.004	0.758
Child Day Care Services	0.156	0.158	0.277	0.072	-0.155	0.002	0.681
Commercial Banking	0.09	0.943	0.527	0.674	-0.389	0	0.779
Computer Systems Design and Related Services	-0.244	0.302	0.46	0.109	0.306	0.005	0.704
Continuing Care Retirement Communities	0.069	0.854	-0.187	-0.103	-0.003	0.005	0.359
Dentist Offices	0.381	-0.891	0.045	-0.022	-0.29	0	0.825
Elementary and Secondary Schools	0.317	-0.558	-0.222	0.023	0.016	0.007	0.8
Fitness and Recreational Sports Centers	-0.949	0.524	-0.181	0.355	1.937	-0.006	0.273
Full-Service Restaurants	0.246	-0.318	0.493	0.053	-0.015	0.003	0.896
Hair, Nail, and Skin Care Services	0.153	0.162	0.488	0.055	0.182	-0.001	0.816
Home Health Care Services	-0.33	0.491	0.435	-0.011	0.365	0.001	0.575
Hotels (except Casino Hotels) and Motel	0.159	0.562	0.676	-0.471	0.798	-0.002	0.636
Insurance Agencies and Brokerages	0.798	-0.061	-0.033	0.183	-0.236	-0.003	0.639
Janitorial Services	0.571	-0.117	-0.195	0.226	0.252	0.003	0.839
Lawyer Offices	0.028	0.19	0.13	0.139	0.109	0.002	0.685
Limited-Service Restaurants	0.461	-0.088	0.433	0.214	-0.096	-0.002	0.846
Management Consulting Services	0.104	0.12	0.565	0.058	-0.098	0.002	0.594
Nursing Care Facilities	-0.191	0.157	-0.02	0.008	-0.294	0.013	0.494
Other Scientific and Technical Consulting	-0.321	0.714	0.422	-0.066	-0.892	0.015	0.594
Pharmacies and Drug Stores	0.236	0.148	0.227	0.035	-0.076	0	0.635
Physician Offices	0.035	0.609	0.332	0.406	0.422	-0.005	0.776
Residential Building Construction	0.086	0.786	0.533	-0.056	0.296	0.003	0.636
Supermarkets and Other Grocery	0.108	0.299	0.422	-0.002	-0.029	0.001	0.74
Veterinary Services	0.46	-0.607	0.489	0.37	-0.49	-0.003	0.856

* The increment is the ratio of Pasadena MW to the California state MW
Green or Red: This result is individually significant

increases in the California state minimum wage because the California minimum wage rose by $4 from $8 per hour in 2011 to $12 per hour in 2019, while the Pasadena minimum wage has risen above the California minimum wage by 50 cents in the second half of 2016, and by $1.50 during the second half of 2017 and by $2.25 in the second half of 2018 and the second half of 2019.

We find significant impact of the rising California state minimum wage on earnings per quarter for many industries. We have highlighted four industries because they form a relatively large part of the Pasadena labor force, they have a high proportion of workers working within \$2 of the minimum wage, and our model specification suggests that the rise in minimum wages has a positive impact on earnings: full and limited service restaurants, supermarkets, and hair, nail, and skin care services.

Table 2.6: Regression Result for Predicting Impact of Minimum Wage on Employment

Industry	log(MW)	log(Incre*)	$y_{(t-1)}$	$log(Y_t)$	$log(Y_{t-1})$	Quarter	R^2
Accounting, Tax Preparation, Bookkeeping	-0.263	0.413	0.803	0.338	-0.109	0.003	0.977
Child Day Care Services	-0.017	-0.055	0.818	1.016	-1.092	0.002	0.948
Commercial Banking	0.079	0.082	0.813	-0.621	0.131	0.003	0.975
Computer Systems Design and Related Services	-0.134	-0.021	0.927	0.464	0.009	0	0.967
Continuing Care Retirement Communities	-0.391	1.071	0.865	-0.85	-0.563	0.014	0.977
Dentist Offices	-0.032	0.124	0.879	0.299	-0.068	0	0.988
Elementary and Secondary Schools	-0.118	0.6	0.893	0.67	-0.808	0.001	0.975
Fitness and Recreational Sports Centers	0.044	-0.095	0.285	-1.004	-2.509	0.051	0.977
Full-Service Restaurants	-0.129	0.157	0.819	0.313	-0.202	0.002	0.989
Hair, Nail, and Skin Care Services	-0.039	-0.126	0.847	0.148	0.224	-0.001	0.955
Home Health Care Services	-0.625	1.089	0.809	0.573	-1.081	0.008	0.978
Hotels (except Casino Hotels) and Motel	-0.01	0.069	0.733	0.763	-0.486	-0.001	0.981
Insurance Agencies and Brokerages	-0.045	0.295	0.827	0.293	-0.315	0.002	0.962
Janitorial Services	-0.686	-0.247	0.681	4.938	-4.16	0.007	0.994
Lawyer Offices	-0.102	-0.051	0.895	-0.024	-0.02	0.002	0.991
Limited-Service Restaurants	-0.005	-0.397	0.753	0.508	-0.483	0.004	0.986
Management Consulting Services	0.445	-0.359	0.764	0.801	-0.71	-0.004	0.921
Nursing Care Facilities	0.062	0.271	0.825	0.247	-0.298	-0.003	0.986
Other Scientific and Technical Consulting	-0.807	1.099	0.85	-0.648	0.377	0.005	0.89
Pharmacies and Drug Stores	-0.136	0.232	0.902	0.321	-0.391	0.002	0.946
Physician Offices	-0.281	0.551	0.724	0.752	-0.672	0.004	0.981
Residential Building Construction	-0.073	0.324	0.842	-0.324	-0.373	0.006	0.94
Supermarkets and Other Grocery	-0.06	0.02	0.76	0.319	-0.228	0.001	0.955
Veterinary Services	-0.306	0.384	0.857	2.31	-1.939	0.005	0.972

* The increment is the ratio of Pasadena MW to the California state MW

Green or Red: This result is individually significant

Table 2.7: Regression Result for Predicting Impact of Minimum Wage on Establishments

Industry	log(MW)	log(Incre*)	$y_{(t-1)}$	$log(Y_t)$	$log(Y_{t-1})$	Quarter	R^2
Accounting, Tax Preparation, Bookkeeping	-0.342	0.449	0.844	-0.112	0.512	0.004	0.991
Child Day Care Services	-0.018	-0.023	0.906	-0.335	0.408	0.001	0.963
Commercial Banking	0.165	-0.461	0.746	1.392	-0.191	-0.006	0.962
Computer Systems Design and Related Services	-0.15	0.503	0.851	0.322	-0.001	0.001	0.968
Continuing Care Retirement Communities	-0.497	1.337	0.934	1.006	0.185	-0.001	0.857
Dentist Offices	-0.11	0.135	0.888	-0.086	0.074	0.002	0.994
Elementary and Secondary Schools	-0.193	0.523	0.881	0.178	-0.205	0.001	0.979
Fitness and Recreational Sports Centers	0.108	0.105	0.812	0.009	0.63	0	0.935
Full-Service Restaurants	-0.061	0.04	0.838	0.089	-0.046	0.001	0.976
Hair, Nail, and Skin Care Services	-0.373	0.23	0.845	-0.586	0.577	0.007	0.974
Home Health Care Services	-0.683	0.322	0.597	-0.289	0.4	0.009	0.967
Hotels (except Casino Hotels) and Motel	-0.23	0.335	0.745	-0.924	1.188	0.004	0.978
Insurance Agencies and Brokerages	0.044	-0.026	0.889	0.17	0.036	-0.001	0.989
Janitorial Services	-0.529	0.142	0.767	-0.049	0.15	0.01	0.957
Lawyer Offices	-0.232	0.232	0.846	0.357	-0.115	0.003	0.997
Limited-Service Restaurants	-0.106	-0.091	0.823	-0.14	0.193	0.003	0.988
Management Consulting Services	-0.649	0.98	0.777	0.831	-0.047	0.01	0.975
Nursing Care Facilities	0.133	-0.109	0.891	0.487	-0.436	-0.004	0.94
Other Scientific and Technical Consulting	-0.609	0.645	0.843	0.152	-0.7	0.007	0.976
Pharmacies and Drug Stores	0.068	0.112	0.786	0.063	-0.015	0.001	0.951
Physician Offices	-0.258	0.302	0.73	0.064	0.106	0.004	0.998
Residential Building Construction	-0.212	0.418	0.809	0.377	0.351	0	0.986
Supermarkets and Other Grocery	0.024	-0.063	0.933	-0.582	0.719	-0.001	0.975
Veterinary Services	-0.159	0.226	0.847	0.096	0.221	0.003	0.985

* The increment is the ratio of Pasadena MW to the California state MW
Green or Red: This result is individually significant

Table 2.5,2.6, and 2.7 present regression result of the main specification using earnings per person, employment, and number of firms as dependent variable. Table 2.8 shows long run impact of state minimum wage change and increment change on earnings per person, employment, and number of establishment. We found significant positive impact of minimum wage on earnings for both full and limited service restaurants, janitorial services, dentist office, elementary and secondary schools, and insurance agencies and brokerages. Negative coefficient on increment ratio means that if Pasadena increases its minimum wage above state minimum wage, the impact of minimum wage on earning is going to be less positive than increase the state minimum wage. We found negative coefficient for increment for continuing care retirement, dentist office, and full-service restaurants. There is less significant impact of minimum wage on employment and number of establishments. The coefficient on lag term is also different across three variables. We can see that earnings per person has least significant coefficient on lag terms compared with the coefficient of lag term on employment and number of firms. This is consistent with our expectation since number of employment

Table 2.8: Regression Result for Long Run Impact of Minimum Wage on Earnings per person, Employment, and Number of Establishments

	Earnings per person		Employment		Establishments	
Industry	LR MW	LR Incre	LR MW	LR Incre	LR MW	LR Incre
Accounting, Tax Preparation, Bookkeeping	0.404	-0.921	-1.336	2.099	-2.195	2.882
Child Day Care Services	0.215	-0.219	-0.0957	-0.3	-0.187	-0.251
Commercial Banking	0.191	1.995	0.423	0.441	0.651	-1.817
Computer Systems Design and Related Services	-0.451	-0.56	-1.82	-0.283	-1.012	3.385
Continuing Care Retirement Communities	0.058	-0.719	-2.884	7.907	-7.541	20.28
Dentist Offices	0.398	-0.932	-0.261	1.027	-0.988	1.207
Elementary and Secondary Schools	0.26	-0.457	-1.098	5.605	-1.624	4.414
Fitness and Recreational Sports Centers	-0.803	0.443	0.0609	-0.133	0.576	0.559
Full-Service Restaurants	0.486	-0.627	-0.71	0.867	-0.373	0.248
Hair, Nail, and Skin Care Services	0.299	-0.317	-0.253	-0.822	-2.4	1.482
Home Health Care Services	-0.584	0.87	-3.273	5.709	-1.694	0.799
Hotels (except Casino Hotels) and Motel	0.491	-1.733	-0.0372	0.258	-0.899	1.311
Insurance Agencies and Brokerages	0.773	-0.0587	-0.261	1.701	0.392	-0.234
Janitorial Services	0.477	-0.0982	-2.152	-0.774	-2.269	0.607
Lawyer Offices	0.032	0.218	-0.97	-0.487	-1.509	1.506
Limited-Service Restaurants	0.812	-0.156	-0.0192	-1.606	-0.6	-0.517
Management Consulting Services	0.24	0.275	1.882	-1.52	-2.906	4.386
Nursing Care Facilities	-0.188	-0.154	0.354	1.548	1.221	-1.002
Other Scientific and Technical Consulting	-0.556	1.235	-5.391	7.34	-3.872	4.1
Pharmacies and Drug Stores	0.306	0.192	-1.387	2.361	0.319	0.523
Physician Offices	0.0531	-0.912	-1.016	1.994	-0.955	1.12
Residential Building Construction	0.184	-1.683	-0.462	2.049	-1.109	2.19
Supermarkets and Other Grocery	0.187	-0.518	-0.248	0.0835	0.36	-0.932
Veterinary Services	0.901	-1.188	-2.144	2.691	-1.042	1.479

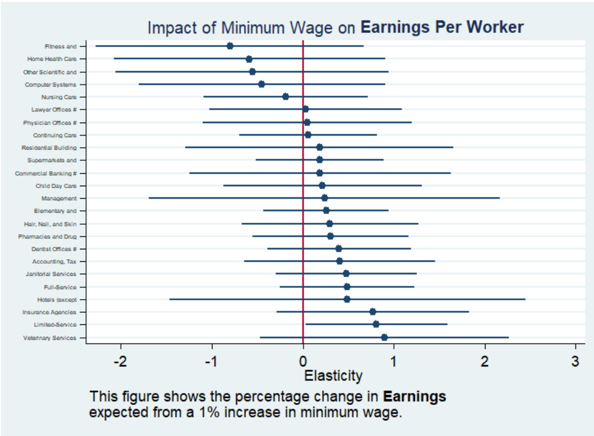
* The increment is the ratio of Pasadena MW to the California state MW
Green or Red: This result is individually significant

and firms need longer time to adjust. Therefore, we expect high correlation across time for these two variables, but not for earnings, since it is easier to adjust.

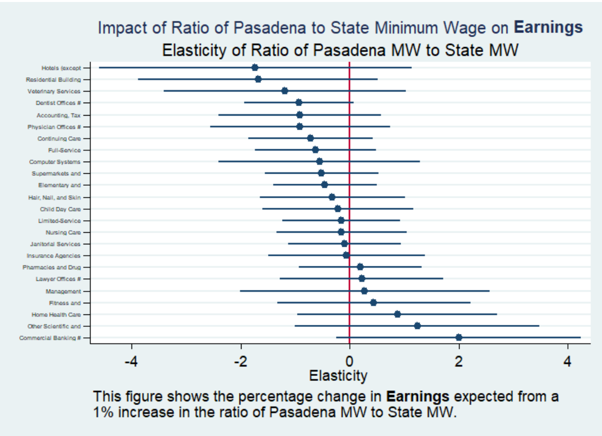
We also display the estimated effects of both the prevailing minimum wage inclusive of any local increment and also effect of the local increment on earnings per worker, number of workers and number of establishments for each industry. In Figure 2.11, each estimate is surrounded with corresponding 95% confidence interval. These estimates are based on the preferred model described in detail above. The estimated impacts below include the estimated impact of minimum wages on all the industries for which we have a complete set of data points over our time period. Many of these industries are not comprised of many minimum wage workers, therefore we would not expect to find a strong impact of minimum wages on these industries at all.

Figure 2.11 (a) presents the impact of minimum wage on earnings per worker by industry. Veterinary Service, Hotels, and Limited-Service restaurants have the largest point estimates. Figure 2.11 (b) shows that differential impact of a local Pasadena increment. A negative estimate indicates a smaller impact of a Pasadena minimum wage increase on earnings. Notice that nearly all of our results are not individually statistically significant. Figure 2.11 (c) presents the impact of minimum wage on employment by industry. Other scientific and technical consulting, Home health care, and Continuing care have the largest negative point estimates. The negative impact of minimum wage on the two consulting industries are quite surprising because they are do not have a large proportion of their workforce working at minimum wage. Our analysis shows that these two industries have been shrinking nationwide as well. Furthermore, the broader category of consulting firms in general (NAICS code 541) has remained stable over this time period. Therefore there is evidence that the decrease is due to the reclassification of many firms in the "Other Technical Consulting" sector to a different consulting designation. Figure 2.11 (d) shows that differential impact of a local Pasadena increment. A negative estimate indicates a stronger negative impact of a Pasadena minimum wage increase on employment. Figure 2.11 (e) presents the impact of minimum wage on

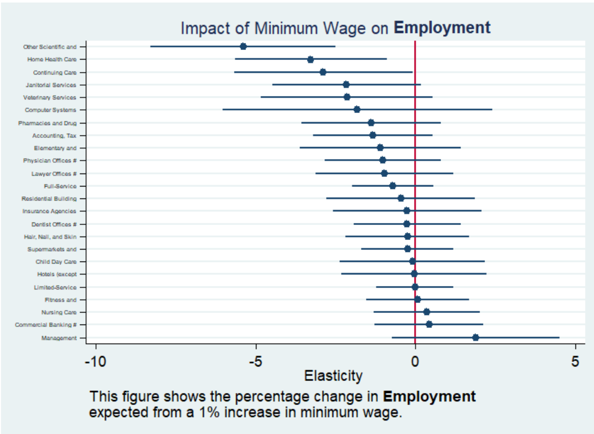
Figure 2.11: Impact of Minimum Wage on Earnings per person, Employment, and Number of Establishments



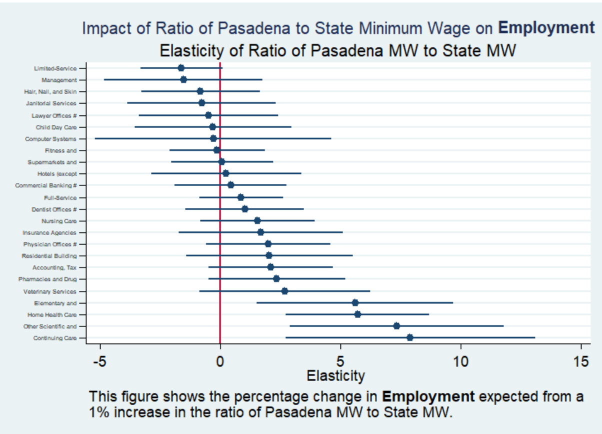
(a)



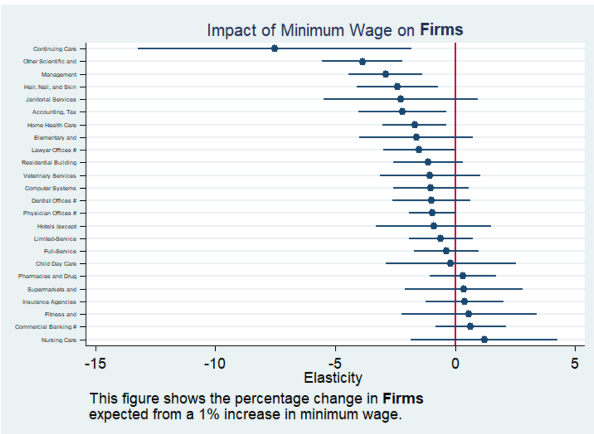
(b)



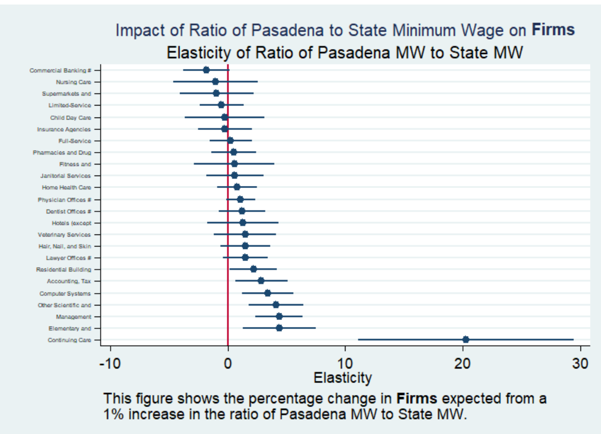
(c)



(d)



(e)



(f)

establishments by industry. Continuing care, Other scientific and technical consulting, and Management Consulting have the largest negative point estimates. Nursing and continuing care also exhibit negative establishment effects. Figure 2.11 (f) shows that differential impact of a local Pasadena increment. A negative estimate indicates a stronger negative impact of a Pasadena minimum wage increase on establishments.

2.5.2 Impact of Minimum Wage on Sales Tax Revenue

Sales tax revenue data that have been provided to us by the city of Pasadena can also be explored for minimum wage effects. Although this dataset does not break Pasadena and the surrounding regions apart into smaller pieces like the QCEW data, it does include data from nearby other cities that are similar to Pasadena in terms of income. These other cities are: Glendale, Monrovia, Santa Monica, and West Hollywood.

Figure 2.12 illustrates the increasing importance of food services as a source of tax revenue for Pasadena since 2011. Casual dining, quick-service dining, and fast casual dining have all experienced substantial increases in tax revenue since 2011 while apparel and specialty

Figure 2.12: Pasadena Sales Revenue

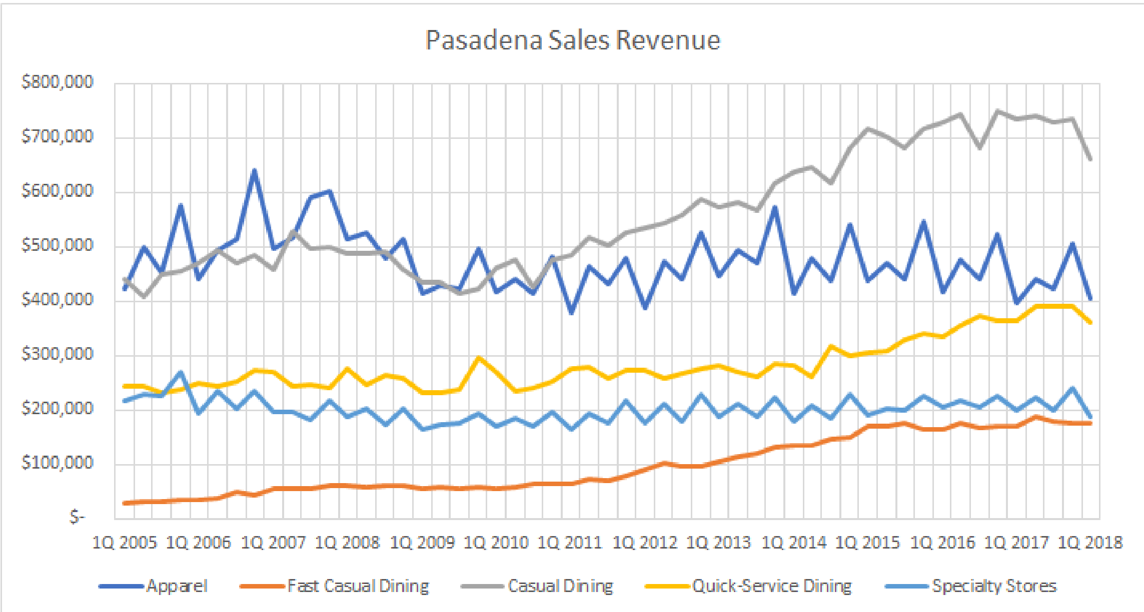
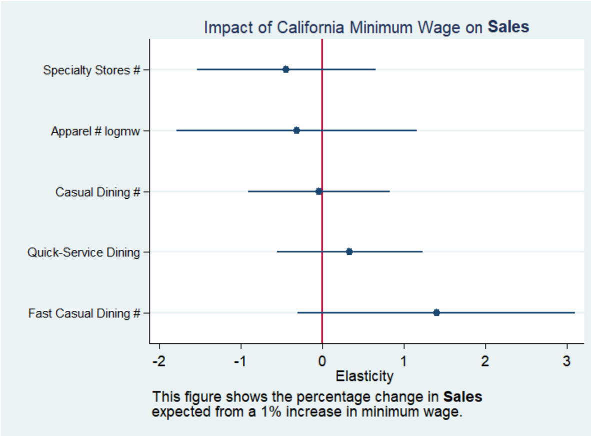
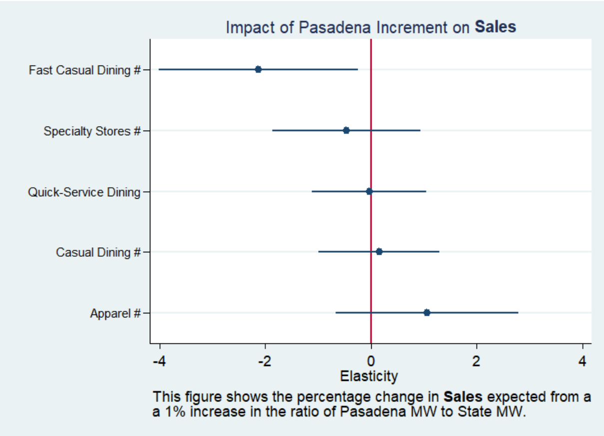


Figure 2.13: Impact of Minimum Wage on Sales Tax Revenue



(a)



(b)

stores have been quite stable.

The timing of the rise in tax revenue from the restaurant sectors after 2011 suggests that the tax revenue is favorably affected by the rise in the minimum wage. A positive impact of minimum wages on sales revenue can occur either because more quantity is sold or because prices rise. A reason why more quantity might be sold is that the added income of restaurant workers allows them to buy more of their own product. A more likely story is that the increase in minimum wages is passed on to customers via higher prices. And of course there may be reasons for increases in price or increases in sales volumes that have nothing to do with the minimum wage.

We can use the same specifications as we have in our previous analysis of earnings, employment, and establishments to examine the impact of minimum wage on sales revenue in these five industries. Figure 2.13 (a) shows the impact of an increase in minimum wages inclusive of the local increment. The solid dots are our point estimates, which show that for Fast Casual Dining, (for example: McDonalds), a 1% increase in minimum wage would result in a 1% increase in sales revenue. The line intervals indicate a 95% confidence interval of our point estimates, and if the lines intersect the solid red line at 0, then our point estimates are not statistically significant at the 5% confidence level. We can see that none of our point estimates of the impact of minimum wage on sales revenue is statistically significant, although we could say that cheaper restaurants seem to have a stronger response than more expensive restaurants and clothing stores. For the restaurants classified as fast casual dining, the evidence says that we would see only 30% of the increase in sales revenue in response to minimum wages would occur within three months.

Figure 2.13 (b) illustrates the separate impact of the Pasadena increment to the minimum wage. Our model includes two minimum wage variables, one is the prevailing minimum wage inclusive of the Pasadena increment and the other is the Pasadena increment separately. If the Pasadena increment behaves just like the California increment, then this second variable would have a zero effect. Once again we would like to stress that we do not have much evidence of this second effect because the Pasadena minimum wage has only risen above the California minimum wage briefly three times in our dataset (which spans to the 2018q1). From the line intervals displayed we can see that only the fast casual dining effect is bounded away from zero, suggesting that the Pasadena increment has much less of an impact than a statewide increase in minimum wages.

Overall our evidence says that sales revenue has a stronger response to minimum wages for restaurants that are cheaper and faster, while restaurants that are more expensive, clothing, and specialty stores do not show evidence of a response.

2.5.3 Impact by January 2021

This part provides a forecast of what happens to earnings, employment, and firms in four specific industries through January 2021 under two different choices for Pasadena's minimum wage, $15 if Pasadena opts for the $1 local increment dictated by the higher minimum wage schedule of the City of Los Angeles (Scenario 1), or $14 if the California minimum wage is used (Scenario 2). (In early February 2019. Pasadena voted in favor of continuing on the same minimum wage schedule as the City of Los Angeles. Therefore we will see if our predictions bear out.)

Table 2.9: Estimated Impact of Minimum Wage by January 2021

Industry	Average Earnings per Quarter	Scenario 1 Potential Increase	Scenario 1 Percent Increase	Scenario 2 Potential Increase	Scenario 2 Percent Increase
Full-Service Restaurants	6379	485	7.60%	553	8.67%
Limited-Service Restaurants	5298	583	11.00%	322	6.07%
Supermarkets and Groceries	7385	255	3.45%	439	5.95%
Hair, Nail, and Skin Care Services	5240	238	4.55%	246	4.69%

Industry	Average Total Employment	Scenario 1 Jobs at risk	Scenario 1 Percent at risk	Scenario 2 Jobs at risk	Scenario 2 Percent at risk
Full-Service Restaurants	6361	-700	-11.00%	-776	-12.20%
Limited-Service Restaurants	4662	130	2.78%	699	15.00%
Supermarkets and Groceries	2488	-85	-3.43%	-54	-2.19%
Hair, Nail, and Skin Care Services	539	-10	-1.79%	34	6.32%

Industry	Average Total Firms	Scenario 1 Firms at risk	Scenario 1 Percent at risk	Scenario 2 Firms at risk	Scenario 2 Percent at risk
Full-Service Restaurants	257	-14	-5.40%	-11	-4.45%
Limited-Service Restaurants	235	-16	-6.95%	4	1.49%
Supermarkets and Groceries	42	3	6.52%	5	10.80%
Hair, Nail, and Skin Care Services	105	-36	-34.50%	-29	-27.60%

For Scenario 1, in January 2021, the Pasadena will have a prevailing minimum wage of \$15, equal to the California level of \$14 plus the \$1.00 Pasadena increment. This involves a rise in the prevailing minimum wage from \$13.25 to \$15 and a fall in the Pasadena increment from \$1.25 to \$1.00. Table 2.9 shows the forecasted long run impact of this rise from 13.25 to \$15 with the local increment equal to \$1. For earnings, we expect to see 4% to 6% increase. We also expect some negative impact on employment and number of establishments, especially in full-service restaurants and hair, nail, and skin care services.

In Scenario 2, the California state minimum wage will increase to \$14 by January 2021.

Table 2.9 report the estimated impact of \$13.25 to \$14 (+5.7%) while dropping the Pasadena increment from \$1.25 to \$0.00 (-7.5%). This is predicted to increase average quarterly earnings in limited-service restaurants by 6.07% and full service restaurants by 8.67%, and to reduce the number of hair, nail and skin salons by 27.60%.

2.6 Robustness Check

2.6.1 Without Pasadena Increment

Adding an additional term representing Pasadena increment does not change the impact of minimum wage on earnings, employment, or number of establishments, except for limited-service restaurants. Adding the local increment of minimum wage would help us understand a local Pasadena minimum wage differs from a statewide minimum wage. The negative impact of Pasadena increment on employment level for limited-service restaurants shows evidence that when local minimum wage increase, minimum wage jobs may migrate to nearby areas.

2.6.2 Without Time Trend

When time trend is not included in the specification, we observe more industries with statistically significant impact from minimum wage on earnings. This is because both earnings and minimum wages are generally increasing over time. Even without increasing minimum wages, historically we observe earnings increase over time due to inflation. Without controlling for time trend, we would mix the increase of earnings due to inflation with the impact of minimum wage.

There is little evidence of impact of minimum wage on employment with or without the time trend. More industries have significant negative impact of minimum wage on establishments when time trends are added. As the figure to the right shows, some industries exhibit increasing establishments until minimum wages are increased. Therefore adding a time trend allows us to project the number of establishments that would have been there had there not been a minimum wage increase.

2.6.3 Analysis for Groups Separately

The results we have discussed so far are using all the groups (The groups are separated by income level. Group 1 has the least income, and Group 5 has the most). We also conduct analysis for each group separately. The purpose of doing group-wise analysis is to examine whether a change in minimum wage has different impact depending on the income level of the affected area. We find that the impact of minimum wage differs little across groups but there is no obvious pattern. Groups 2 and 3 provide the most significant evidence that increase in minimum wage would increase earnings. Group 4 presents the weakest evidence. Most industries show different results across different groups. However, for Full and Limited service restaurants, there are consistent results across all groups showing that an increase in minimum wages would increase earnings. There is little evidence of the impact of minimum wages on employment or number of establishments.

2.7 Conclusion

We have used the data available to us to analyze the impact of minimum wages on earnings, employment, establishments, and sales tax revenue. We find that minimum wages have a measurable impact on earnings for low wage industries (such as full and limited service restaurants), and our preferred model shows a significant negative impact of minimum wages on the number of Hair/Nail Salons and also a negative impact of the Pasadena increment on the number of jobs in Limited Service Restaurants. However, we do obtain negative estimates of the impact of minimum wages on employment and establishments in most industries.

This study has difficulty detecting the impact of minimum wages on employment and establishments because firms may anticipate upcoming changes in minimum wages, and also may adjust everything but wages slowly over time. Indeed our own estimates show that only one-fifth of the impact of an increase in minimum wages would show in the data within three months. Data from the rest of 2018 would be quite helpful because 2018 and 2019 are the years during which the Pasadena minimum wage is highest above the California state minimum wage.

We find evidence that 50% of the impact of minimum wages on earnings is realized in the first quarter, while only 20% of the impact of minimum wages on employment of on establishments is realized in the first quarter.

We find smaller estimates in general of the impact of a Pasadena increment than a Statewide increment, however jobs in limited service restaurants show evidence of leaving Pasadena in higher numbers when the difference between the Pasadena minimum wage and the California minimum wage is greater. An additional year of data and the corresponding greater time and greater difference between the Pasadena and Statewide minimum wage levels would allow us to more accurately estimate the separate effect of the Pasadena increment.

CHAPTER 3

Chinese Environmental Regulation, Its Effect on Economic Activities and Industry Composition

3.1 Introduction

Urbanization, motorization, and industrialization have intensified both growth and pollution levels in developing countries over the past 30 years. Developing countries suffer the most from environmental deterioration, and they generally have the weakest environmental regulations, if they have them at all.

China, the world's largest developing country, has seen dramatic income growth since opening up in 1978. The central government employed a series of reform policies such as low rent and low tax rates to attract foreign investment. It also encouraged private entrepreneurs to start businesses and it privatized many state-owned enterprises. As a result, the Chinese economy has enjoyed an average growth rate of more than 9% per year since then. However, this unprecedented economic growth has come at a price. During the past 30 years, China has also become one of the dirtiest countries in the world. China's domestic environmental issues have drawn unflattering attention from all over the world, and future Chinese pollution levels will have an increasingly large effect on pollution levels worldwide. Meanwhile, as Chinese cities' overall income increases, people have begun demanding a cleaner living environment. With the spread of the Internet, blogs, micro blogs, cell phone apps, and so on, it is ever harder for the Chinese government to hold their monopoly on the information Chinese citizens have access to. This pressures the Chinese government to prioritize environmental protection, rather than simply focus on income growth. Yet, how these policies affect industrial pollution levels, economic activities, and industry composition are still unclear.

In this paper, I seek to answer three questions: one, is environmental regulation in China effective reducing industrial emissions? Two, how do environmental regulations affect economic activities? Three, how has environmental regulation affected industry composition? I am also interested in how the change in political regime has affected the regulations.

To answer these questions, I focus on the Two-Control-Zones (TCZ) policy. A detailed description of this policy follows in Section 3.2. The TCZ policy, enacted in 1998, assigned some cities to one of two control zones. If cities are designated as TCZ (that is, if they fall into one of the two zones), they are obligated to follow more stringent environmental regulations. By using newly constructed ambient pollution data from NASA satellites and city-level industrial SO_2 emissions data from 2003 to 2012, I investigate whether being designated as TCZ improved cities' environmental performance. To study the effect of TCZ on economic activities, I look at city-level GDP growth, foreign direct investment (FDI), population, and industrial output. I use the number of firms and employment in each industry sector in each city from 1998 to 2007 to study the TCZ policy's effect on industry composition. Therefore, I can observe whether the TCZ policy drives polluting industries to move to or open up more in less-regulated areas. I use the Eleventh Five-Year Plan (see Section 3.2) as an indicator of the change in the government's attitude toward environmental protection to study whether the effect of the TCZ policy is different before and after the Eleventh Five-Year Plan.

The main empirical difficulty I face is that TCZ cities were not randomly assigned. Cities were designated as TCZ based on their pollution level, income, population, and other variables. Therefore, TCZ and non-TCZ cities have different characteristics. Therefore, just comparing TCZ and non-TCZ cities yields biased coefficients. To solve this problem, I use propensity-score matching. Based on city-level pollution, income, population, industrial output, and FDI in 1997, I estimate the propensity score as the probability for each city to get designated as TCZ. I match TCZ cities with non-TCZ cities that have the closest propensity score. By comparing the matched cities, I glean the average TCZ treatment effect.

Using ambient aerosol optical thickness data from NASA as an approximate measure of ambient air pollution, I find evidence that TCZ cities had 5% lower ambient pollution in 2000 to 2012 compared with non-TCZ cities. After the Eleventh Five-Year Plan, this difference increased to 7%. Even though I do not find evidence that the TCZ policy causes a greater industrial SO_2 decrease in the 2003–2012 period, I do find evidence that SO_2 emissions increased more in TCZ cities (relative to non-TCZ cities) before 2006 and decreased more in TCZ cities from 2006 to 2010, the period of the Eleventh Five-Year Plan. The reduction in SO_2 emissions in the production process is higher in TCZ cities — this result is robust across different specifications and time periods. In terms of economic activities, the TCZ policy causes 8% lower GDP growth and almost 30% less new FDI in cities designated as TCZ, especially during the Eleventh Five-Year Plan. TCZ cities also attracted 4% less population. Regarding industry composition, I find evidence that TCZ cities have higher increase in number of firms and employment in dirtier industries.

The economic does contain a few studies on the effects of environmental regulation, most of which have concentrated on the Clean Air Act in the United States. Greenstone (2004) studies the impact of the Clean Air Act on sulfur dioxide concentrations. He concludes that the nonattainment designation played only a minor role in the reduction of sulfur dioxide. Henderson (1996) and Becker and Henderson (2000) studied the costs of environmental regulation. Their work recorded evidence that the Clean Air Act has driven polluting industries to open or relocate in areas with less-stringent environmental regulations. On the other side, Chay and Greenstone (2003, 2005) focus on the benefits of environmental regulation. They find evidence that the Clean Air Act is correlated with lower infant mortality, and that a cleaner environmental is correlated with higher local housing prices.

The literature also contains some work on developing countries. Greenstone and Hanna (2014) find some evidence that India's environmental regulations have caused lower pollution concentrations; this paper further correlates the most successful regulation with a decrease in infant mortality. Similarly, Foster et al. (2009) correlate Mexico's environmental regulations with lower infant mortality. To date, little work has been done on environmental regulation in China. Chen et al. (2013) study the unexpected effect of the Huai River heating policy on the environment. However, this policy is not an environmental policy. Tanaka (2015), like this paper, focuses on the TCZ policy. However, due to data limitations, this paper uses pollution concentration data for only 70 cities (there are around 300 prefecture-level cities in China) and the data cover only up to 2000. To extend this line of research, I use industrial emissions data, which is a more direct measure of the effectiveness of the regulation, since the regulation focuses mainly on reducing industrial emissions. In addition, by using data from 2003 to 2012, I am able to evaluate the long-term effect of this policy and its interaction with political regime change.

This paper contributes to the literature in three ways. One, it uses NASA aerosol depth to measure ambient air pollution levels, providing a longer, more accurate analysis. Two, it presents a more comprehensive evaluation of the TCZ policy—not only the effectiveness of this policy but also how this effectiveness changes with shifts in the government's attitude toward environmental protection—than has been done in previous papers. Three, this paper further analyzes the effect of an environmental regulation on economic activities and industry composition, which, to my knowledge, is the first to do so for developing countries. China's experience will shed light on the issue for other developing countries and help to guide them in policy implementation.

The rest of the paper is organized as follows: Section 3.2 offers some background on Chinese environmental regulation. Section 3.3 describes the data and summary statistics. Section 3.4 covers my empirical methodology. Section 3.5 contains my results. Section 3.6 concludes.

3.2 Background on China's Environmental Regulation

3.2.1 The Two-Control-Zones Policy

The Chinese environment has deteriorated dramatically since the 1990s. To control for pollution emissions, the Chinese central government in 1996 issued its first-ever large-scale environmental regulation. This policy aims to control ambient sulfur dioxide and acid rain in heavily polluted areas in China. The two control zones are the acid-rain control zone and the SO_2 control zone. If cities are assigned to one of these two zones, they are obligated to issue more-stringent environmental regulations.

According to TCZ policy, cities exceeding certain standards are assigned to either the acid-rain control zone or the SO_2 control zone.

Cities would be assigned to the acid-rain control zone if they met the following three conditions:

(1) Monitored raining has a $pH \leq 4.5$.

(2) Sulfate depositions are greater than critical load.

(3) SO_2 emission levels are high.

Cities would be assigned to the SO_2 control zone if they met the following three conditions:

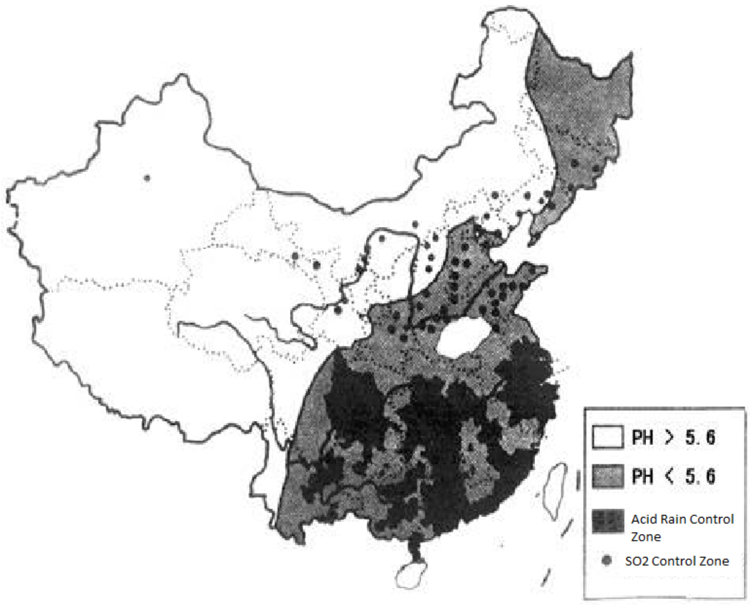
(1) Average annual ambient SO_2 concentration exceeded the Class II standard.

(2) Average daily ambient SO_2 concentration exceeded the Class III standard.

(3) SO_2 emission levels are high.

According to these standard, 64 cities are in the SO_2 control zone and 117 cities are in the acid-rain control zone. In total, 175 cities across 27 provinces are included in the two control zones (see the map in Figure 3.1). As this figure shows, the TCZ are mainly concentrated on the east side of the country, the acid-rain control zone mainly located in the south and the SO_2 control zone in the north. It was hard to set up a single standard that was equally suitable for the southern and northern cities,thus the distinction between the acid-rain and SO_2 control zones. Yet, once cities were designated as TCZ, the regulations implemented were the same. Thus, I am not going to separate these two control zones.

Figure 3.1: Two-Control-Zones



The total TCZ area is around 1.09 million km^2, about 11.4% of Chinese territory. The acid-rain control zones comprise about 800,000 km^2 (8.4% of Chinese territory), the SO_2 control zones comprise about 290,000 km^2 (3% of Chinese territory). Together, these cities contained 40.6% of the national population, generated 62.4% of national GDP, and were responsible for 60% of total SO_2 emissions in 1995.

If cities are designated as TCZ, local governments were required to create plans and set specific targets to reduce sulfur dioxide concentration, set up programs to reduce emissions, and set aside special funding to clean up pollution. To protect the environment, the central government decided to gradually phase out mining of coal containing 3% or more sulfur. In the two control zones, new construction of coal-fired power stations is prohibited. Existing coal-fired power stations were required to install desulfurization devices by 2010. Cities are required to use economic incentives to control sulfur emissions, like taxing or fining major sulfur emitters. The policy also calls for developing new technologies that reduce sulfur emissions.

3.2.2 The Five-Year Plans

Since 1953, the central government has issued a Five-Year Plan every five years. The plan establishes the direction of economic development, setting development targets and launching reforms. In the Tenth Five-Year Plan (2001–2005), the main target was to achieve economic output growth and increase urbanization while controlling population and stabilizing price levels. The Eleventh Five-Year Plan (2006–2010) place greater emphasis on environmental and resource issues.The Eleventh Plan goals in terms of environmental protection and resource conservation were to decrease energy consumption per unit of GDP by 20%, decrease water consumption per unit of industrial added value by 30%, increase the rate of solid industrial waste use from 55.8% in 2005 to 60% in 2010, and decrease total discharge of major pollutants by 10%. The central government distributed these emission-reduction targets to each provincial government. This represents a clear shift in the central government's attitude toward environmental issues. I include the interaction between TCZ and Five-Year Plan because in the Tenth and Eleventh Plans, the central government set up special targets for TCZ cities. Therefore, by signalling that the central government is more serious about environmental regulation, the Eleventh Five-Year Plan may improve the effectiveness of the regulation.

3.3 Data and Summary Statistics

In this section, I describe my datasets and provide the summary statistics.

3.3.1 Data Sources and Description

My four main variables are regulation, environment, economic activity, and industry.

Regulation data focus on the list of cities designated as TCZ. This list is available on the website of the Central People's Government of the People's Republic of China[1]. Of the 281 cities in my dataset, 151 were designated as TCZ.

Environment data include aerosol optical depth, city-level industrial SO_2 emissions, SO_2 reduction, and pollution intensity for each industry sector. The aerosol optical depth data were generated by MODIS [Moderate Resolution Imaging Spectroradiometer]; I acquired

[1] $http://www.law-lib.com/law/law_view.asp?id=108674$

them from NASA's Goddard Earth Sciences Distributed Active Archive Center(GES DAAS). Several studies e.g. Gupta et al.(2006), Chu et al. (2003), Kumar et al. (2007) have shown a promising relationship between ambient air pollution and aerosol optical depth. Data on industry SO_2 emissions, and SO_2 reduction come from the 2003-2012 Chinese Urban Statistics Yearbooks. Data on pollution intensity are not direct available: to calculate pollution intensity, I divide total SO_2 emissions from a certain industry sector by total value added or total employment in this sector. Data on industry-sector-level SO_2 emissions come from the 2001 Chinese Environmental Statistical Yearbook and the sector-level value added and employment data come from the 2001 Chinese Industry Statistical Yearbook.

Economic activity data include GDP, FDI, industry output, and population. These data are all available from the 2003–2012 Chinese Urban Statistics Yearbooks. City-level industry-sector data are aggregated from firm-level data. The firm survey data come from a nonpublic source; they cover firms that made over 5 million in revenue per year from 1998 to 2007. The year-to-year variables are not entirely consistent over this period; however, the information for firms' location and number of employees are always available. Therefore, I use number of firms and number of employees as the main measurements.

One problem is that all the data I introduce above are from the post-TCZ in the period. Therefore, I also need information for the earlier period to create propensity scores. The prefecture city-level industrial emissions in SO_2 and control variables include GDP and FDI, and industry data come from the 1997 Urban Statistics Yearbook. Of the 223 prefecture cities included in this data, 140 were designated as TCZ.

3.3.2 Summary Statistics

In this section, I present summary statistics for the main variables. In Table 3.1, I present 1997 summary statistics for environment, area, population, GDP, industry, and FDI. A clear pattern emerges from this table: TCZ cities are more polluted, more populated, have higher GDP, and attracted more FDI yhsn non-TCZ cities in 1997, the year before the TCZ policy

Table 3.1: Summary Statistics: Non-TCZ and TCZ Cities, 1997

	Non-TCZ Cities				
Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Environment:					
SO_2 Emissions (thousand tons)	70	316.0905	498.1192	2.124	2655.38
SO_2(tons)/km^2	70	31.61429	46.86004	1	260
Area and Population:					
Area(km)	83	14362.64	11915.92	236	68726
Population (1000)	83	352.6357	251.1626	13.09	1360.04
GDP:					
GDP (million)	82	1907.429	1460.302	56.844	7462.006
Urban GDP (million)	83	747.9753	865.5728	56.844	5395.539
GDP per capita (million)	82	6.414313	5.385234	1.880632	37.21254
Industry:					
Industry Output (1,000)	83	2653.339	2362.361	73.955	9998.064
Urban Industry Output (million)	83	1056.369	1205.571	52.518	7285.255
Foreign Direct Investment:					
FDI Contract ($ 1,000)	78	5008.115	7873.94	14	46734
Urban FDI Contract (1,000)	73	3244.589	5500.215	20	38245
Actual FDI ($ 1,000)	77	4534.987	6414.101	14	29572
Urban Actual FDI ($ 1,000)	72	3099.708	4783.25	8	29572
	TCZ Cities				
	Obs.	Mean	Std. Dev.	Min.	Max.
Environment:					
SO_2 Emissions (thousand tons)	114	795.4638	1352.444	7.83	8502.966
SO_2(tons)/km^2	114	70.52632	94.62064	1	681
Area and Population:					
Area(km)	140	11764.13	11247.28	1113	90021
Population (1000)	140	432.1654	329.9061	38.29	3042.92
GDP:					
GDP (million)	140	3794.463	4113.081	261.888	33602.1
Urban GDP (million)	140	1910.306	3073.119	75.203	26994.7
GDP per capita (million)	140	9.259076	9.666168	1.90105	103.2346
Industry:					
Industry Output (1,000)	140	6004.138	7036.169	323.169	56499.3
Urban Industry Output (million)	140	2951.513	4531.004	80.878	39427.8
Foreign Direct Investment:					
FDI Contract ($ 1,000)	136	33995.54	77641	3	531999
Urban FDI Contract (1,000)	134	21948.28	54643.66	6	466711
Actual FDI ($ 1,000)	135	29530.59	60538.01	3	480816
Urban Actual FDI ($ 1,000)	131	19066.99	49959.01	22	480816

was enacted. This indicates that TCZ status is not a random assignment, which may lead to biased estimates when estimating the effect of this regulation. I use propensity-score matching to solve this issue (see Section 3.4).

Figure 3.2: Industrial SO_2 Emission Trend, 2003–2012

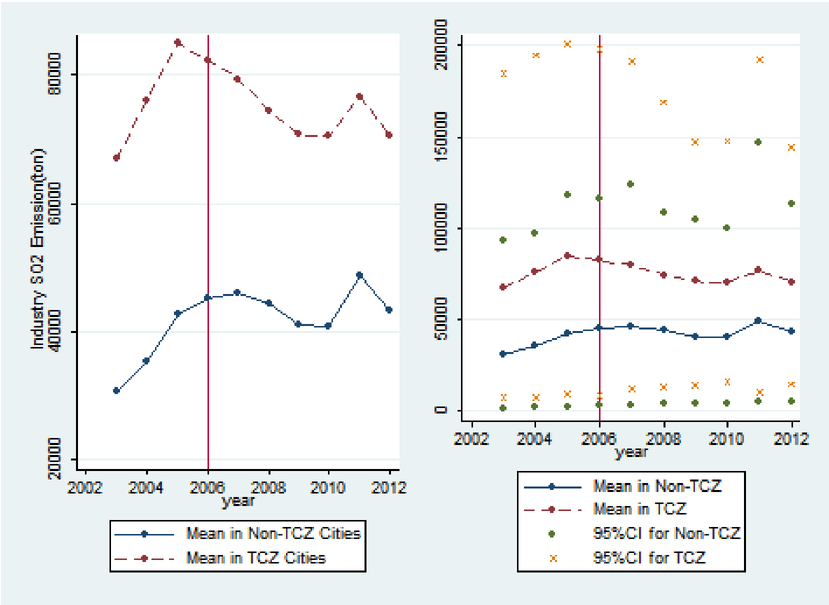
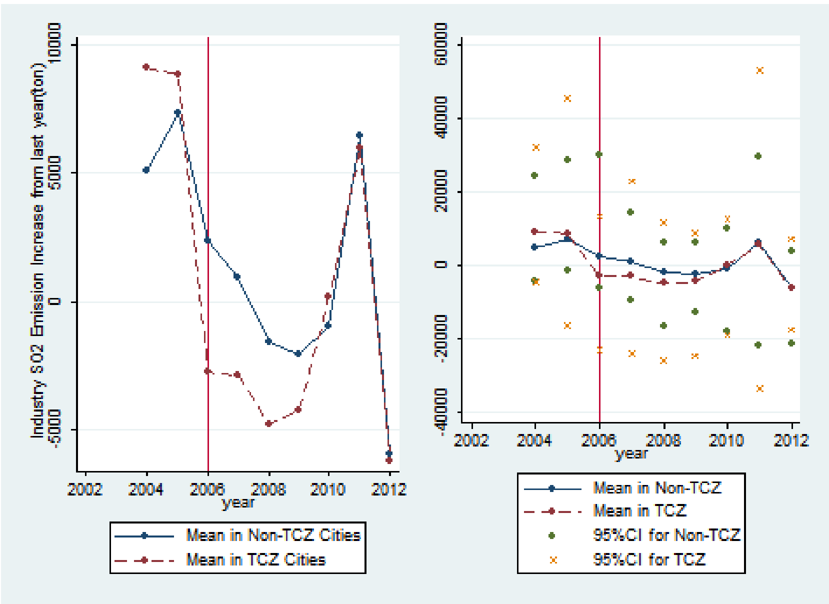


Figure 3.3: Industry SO_2 Emission Increase from Previous Year, 2003–2012



Figures 3.2 through 3.6 present trends for industry emissions, GDP, FDI, population, and industrial output by TCZ and non-TCZ cities from 2003 to 2012. Left-hand-side panels present trends for the mean, and right-hand-side panels also include 95% confidence intervals. Trends for industrial SO_2 emissions from 2003 to 2012 are presented in Figure 3.2. Overall, TCZ and non-TCZ cities exhibit an inverse U-shape pattern though this pattern is more pronounced for TCZ cities. SO_2 emissions start to decrease after 2006 for TCZ cities and after 2007 for non-TCZ cities. The large decrease in industrial SO_2 emissions is driven by the clear decrease in emissions levels in the most polluted TCZ cities,

as shown in the right-hand-side graph. The previous year's industrial SO_2 emissions increase from 2004 to 2012 is illustrated in Figure 3.3. From 2006 to 2009, TCZ cities experienced a larger decrease in mean emissions than the mean decrease in non-TCZ cities. One might expect that this decrease is caused by the economic depression during this period. However, from Figure 3.4, which depicts the GDP per capita growth trend, there is no clear break during this period. Therefore, it is unlikely that the pollution decrease from 2006 to 2009 was caused by the economic showdown. Rather, this pattern is more likely driven by the Eleventh Five-Year Plan, in which the central government prioritized environmental protection and set up more stringent targets for TCZ cities.

Figure 3.4: GDP per Capita Trend, 2003–2012

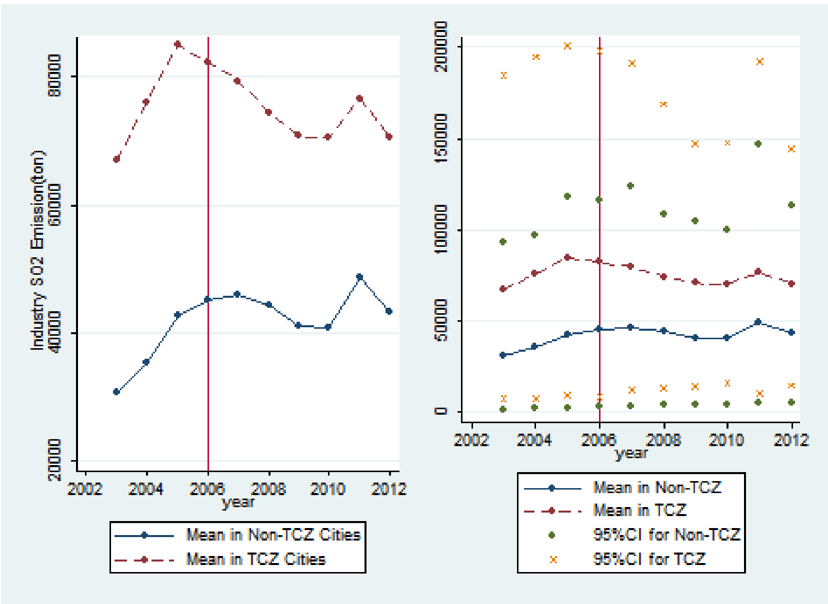
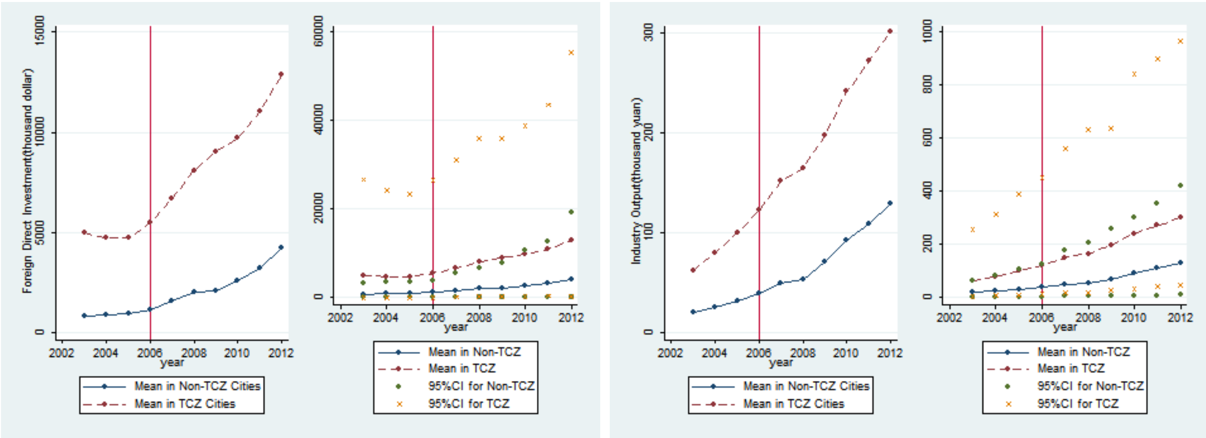


Figure 3.4 depicts that GDP per capita increased dramatically for both TCZ and non-TCZ cities. This is consistent with the fast-paced economic development in China during this period. The right-hand-side panel of Figure 3.4 also depicts the 95% interval for GDP per capita. At the 95% confidence interval of GDP per capita TCZ cities and non-TCZ cities converge, while they diverge at the 5% confidence interval. This indicates that poor TCZ cities developed faster than poor non-TCZ cities, while rich non-TCZ cities developed faster than rich TCZ cities.

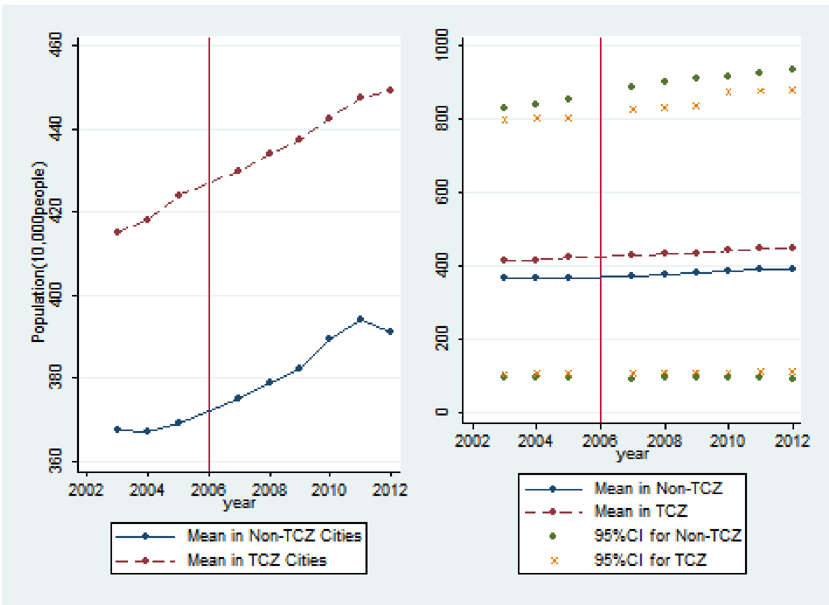
Trends for FDI and industrial output from 2003 to 2012 are depicted in Figure 3.5. TCZ cities grew more both in terms of FDI and industrial output than non-TCZ cities especially after 2006. This larger increase is largely attributed to the growth in 95% CI cities. The

Figure 3.5: Foreign Direct Investment and Industry Output Trend, 2003–2012



population trend during this period, on the other hand, is essentially the same between TCZ and non-TCZ cities (Figure 3.6).

Figure 3.6: Population Trend, 2003–2012



The summary statistics for the above variables for TCZ and non-TCZ cities are depicted in Table 3.2. From 2003 to 2012, TCZ cities had greater aerosol optical depth, generally emitted more SO_2, had a greater reduction of SO_2, and recorded lower increases in SO_2 emissions. TCZ cities also had higher population growth than non-TCZ cities, but this growth was almost equal in these two sets of cities. TCZ cities had higher industrial output and higher industrial output increases. TCZ cities also dominated non-TCZ cities in terms of GDP and FDI.

To study whether the TCZ policy unexpectedly drove polluting industries to migrate

Table 3.2: Summary Statistics, 2003–2012

	Non-TCZ				
Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Pollution:					
Average Aerosol Depth	1024	0.44	0.19	0.07	0.85
SO_2 Emissions (thousand tons)	812	41.96	33.32	0.012	404.34
Amount of SO_2 Reduction (tons)	617	28.53049	71.86957	0	639.559
Difference in SO_2 Emission from last year(tons)	728	1125.827	19361.97	-291063	310075
Population					
City Population (1000)	812	379.8302	247.6477	16.37	1206.3
Population Growth (1/1000)	721	5.658252	4.58532	-6.2	40.78
Industry					
Industry Output (billion)	799	92.31001	115.9786	0.93437	1026.748
Difference in Industry Output from last year (billion)	698	19.05173	25.75252	-71.59761	216.859
GDP					
GDP (million)	799	71374.85	66982.74	3177.31	4.55×10^5
GDP per capita (1,000)	805	23.21748	20.5168	0.099	145.395
GDP growth(%)	804	13.67542	3.313727	-1.2	29.1
FDI					
New Contract	781	46.61972	79.25803	0	915
New Investment ($ million)	795	21.13949	33.71458	0	368.236
	TCZ				
	Obs.	Mean	Std. Dev.	Min.	Max.
Pollution:					
Average Aerosol Depth	1636	0.54	0.18	0.07	1
SO_2 Emissions (thousand tons)	1341	78.37	55.25	0.924	496.377
Amount of SO_2 Reduction (tons)	1069	118.935	720.3439	0	18399.5
Difference in SO_2 Emission from last year(tons)	1200	268.3767	23760.9	-304452	349091
Population					
City Population (1000)	1349	435.9149	235.0025	41.6	1173.3
Population Growth (1/1000)	1203	4.821613	4.123021	-8.9	28.58
Industry					
Industry Output (billion)	1300	225.6176	315.6203	1.57787	2874.554
Difference in Industry Output from last year (billion)	1124	38.56942	52.74671	-236.4673	436.7171
GDP					
GDP (million)	1315	1.42×10^5	1.66×10^5	4868.5	1.36×10^6
GDP per capita (1,000)	1347	30.1491	21.42404	2.882	125.709
GDP growth(%)	1342	13.82401	3.01977	1	37.69
FDI					
New Contract	1327	180.5795	398.2966	0	5096
New Investment ($ million)	1332	71.49893	126.4788	0	1235.033

to less-regulated areas, I need to examine the change in industry distribution in TCZ and non-TCZ cities from 1998 to 2007. In China, industry is divided into about 40 sectors. In Table 3.3, I rank the sectors by their pollution intensity. There are two measures for

Table 3.3: Pollution Intensity by Industry Sector in 2001

Sector	Poll. Int.* (VA)	Rank (VA)	Poll. Int.* (Emp)	Rank (Emp)
Electronic and Equipment Telecommunications	7.52	1	74.66	5
Instruments, Meters, Cultural and Machinery Office	12.47	2	53.52	3
Garments and Other Fiber Products	15.28	3	44.36	1
Petroleum and Natural Gas Extraction	15.95	4	537.9	17
Production and Supply of Tap Water	16.25	5	58.30	4
Cultural, Educational, and Sports Goods	17.24	6	46.36	2
Tobacco Processing	17.57	7	776.7	24
Electric Equipment and Machinery	22.36	8	136.6	10
Plastic Products	24.15	9	112.3	8
Printing and Record Medium Reproduction	25.37	10	113.2	9
Transport Equipment	30.89	11	170.4	14
Leather, Furs, Down, and Related Products	36.14	12	111.4	7
Furniture Manufacturing	36.20	13	142.7	11
Metal Products	39.99	14	172.7	15
Special-Purpose Equipment	41.84	15	143.5	12
Ordinary Machinery	43.14	16	154.1	13
Medical and Pharmaceutical Products	90.39	17	634.1	20
Logging and Transport of Timber and Bamboo	90.48	18	76.25	6
Food Processing	123.48	19	698	22
Timber Processing, Bamboo, Cane, Fiber Palm, and Straw Products	153.51	20	577	19
Food Manufacturing	159.84	21	802	25
Rubber Products	180.39	22	727	23
Textile Industry	189.24	23	549	18
Nonferrous Metals Mining and Dressing	221.11	24	694	21
Coal Mining and Dressing	254.30	25	473	16
Beverage Manufacturing	287.20	26	1943	28
Nonmetal Minerals Mining and Dressing	387.94	27	942	26
Petroleum Processing and Coking	389.79	28	5815	36
Smelting and Pressing of Ferrous Metals	478.76	29	2938	31
Raw Chemical Materials and Chemical Products	490.97	30	2467	29
Chemical Fiber	574.53	31	3168	32
Production and Supply of Gas	577.52	32	1811	27
Papermaking and Paper Products	897.41	33	3744	33
Ferrous Metals Mining and Dressing	950.05	34	2817	30
Smelting and Pressing of Nonferrous Metals	1035.78	35	5602	35
Nonmetal Minerals Products	1298.49	36	4008	34
Production and Supply of Electric Power, Steam, and Hot Water	2690.87	37	31612	37

* : Pollution intensity is calculated as SO_2 emissions (tons) per value added(VA) or SO_2 emissions (tons) per employee(emp).

pollution intensity: per value added and per employee. The first is calculated by dividing the total SO_2 emissions in a sector by the sector's value added, and the second is calculated by dividing the total SO_2 emissions in a sector by the total number of employees in 2001. The unit is SO_2 emissions per value added and per employee. The higher the rank, the more polluted the industry. The two measures of pollution intensity are generally consistent with each other, with some minor differences. Table 3.4 and Table 3.5 presents the distribution of each industry in TCZ and non-TCZ cities in 1998 and 2007. To approach distribution, I calculate the share of employment for each industry sector in each city in 1998 and 2007. Then I calculate the mean and standard deviation of each industry sector's employment share across cities. The resulting pattern is unclear. I will undertake further statistical analysis to study this in Section 3.4.

3.4 Empirical Methodology

3.4.1 Empirical Problem

To analyze whether this TCZ policy is effective in terms of reducing pollution levels and its side effect on slowing economic activities and changing industry composition, I need to estimate the average treatment effect of this policy. Ideally, if the TCZ assignment is totally random, then just by comparing the change of the interest variable between TCZ cities and non-TCZ cities would give us the result. However, the TCZ assignment is not exogenous. TCZ cities were more polluted, were more populated, had higher income levels, and attracted more FDI in 1997 (Table 3.1). Therefore, by simply comparing the pollution change in TCZ cities versus non-TCZ cities does not give me the average treatment effect. Even if I observe that TCZ cities have higher decreases in pollution, I cannot conclude that this is because of the policy—it could just be that higher-income cities have a higher demand for a cleaner environment. Therefore, the effect of the policy alone cannot be identified. To solve this problem, I need to compare cities that were similar to each other before the TCZ policy

Table 3.4: Employment Share by Industry Sector, 1998

		Non-TCZ		TCZ	
Sector	Rank	Mean	Std. Dev.	Mean	Std. Dev.
Electronic and Telecommunications Equipment	1	2.763	3.029	4.006	3.379
Instruments, Meters, Cultural, and Office Machinery	2	2.349	3.535	3.339	4.948
Garments and Other Fiber Products	3	1.734	2.434	4.576	6.545
Petroleum and Natural Gsa Extraction	4	32.10	33.77	2.657	4.335
Production and Supply of Tap Water	5	1.026	0.809	0.819	0.606
Cultural, Educational and Sports Goods	6	0.607	0.890	1.528	2.379
Tobacco Processing	7	1.343	1.286	0.741	0.825
Electric Equipment and Machinery	8	3.578	3.537	3.526	5.279
Plastic Products	9	1.284	1.450	1.923	2.112
Printing and Record Medium Reproduction	10	0.913	0.700	1.069	0.800
Transport Equipment	11	5.657	10.18	4.510	4.322
Leather, Furs, Down, and Related Products	12	1.673	4.972	2.275	3.964
Furniture Manufacturing	13	0.458	0.660	0.413	0.464
Metal Products	14	1.950	2.791	2.684	2.467
Special-Purpose Equipment	15	3.901	2.777	3.950	3.019
Ordinary Machinery	16	4.122	3.581	5.603	4.114
Medical and Pharmaceutical Products	17	1.687	1.488	1.654	1.442
Logging and Transport of Timber and Bamboo	18	12.57	19.27	1.123	1.516
Food Processing	19	5.024	5.805	2.920	3.376
Timber Processing, Bamboo, Cane, Palm Fiber, and Straw Products	20	1.142	1.618	0.844	1.127
Food Manufacturing	21	1.579	1.832	1.871	1.955
Rubber Products	22	1.223	1.339	1.371	1.784
Textile Industry	23	9.003	7.136	8.887	7.874
Nonferrous Metals Mining and Dressing	24	1.307	1.312	1.956	3.856
Coal Mining and Dressing	25	18.94	25.44	13.82	17.03
Beverage Manufacturing	26	2.455	2.930	2.107	2.634
Nonmetal Minerals Mining and Dressing	27	2.015	3.630	1.513	2.379
Petroleum Processing and Coking	28	2.307	4.687	1.752	3.344
Smelting and Pressing of Ferrous Metals	29	4.735	12.04	6.018	10.20
Raw Chemical Materials and Chemical Products	30	5.991	4.237	7.280	5.239
Chemical Fiber	31	1.387	1.736	1.061	1.015
Production and Supply of Gas	32	0.426	0.431	0.360	0.362
Papermaking and Paper Products	33	2.443	2.375	2.217	1.827
Ferrous Metals Mining and Dressing	34	0.818	1.273	1.607	3.395
Smelting and Pressing of Nonferrous Metals	35	1.394	2.279	3.513	9.921
Nonmetal Mineral Products	36	7.628	5.900	8.797	5.975
Production and Supply of Electric Power, Steam, and Hot Water	37	3.760	3.674	3.931	3.658
Overall		3.505	7.663	3.343	5.600

was implemented, but where only some of these cities got designated as TCZ. By comparing these two groups of cities, I am able to identify the average treatment effect of the TCZ

Table 3.5: Employment Share by Industry Sector, 2007

		Non-TCZ		TCZ	
Sector	Rank	Mean	Std. Dev.	Mean	Std. Dev.
Electronic and Telecommunications Equipment	1	2.984	4.023	5.057	7.770
Instruments, Meters, Cultural, and Office Machinery	2	0.943	1.589	1.154	1.223
Garments and Other Fiber Products	3	2.878	3.496	4.581	4.618
Petroleum and Natural Gsa Extraction	4	24.59	27.39	1.256	2.159
Production and Supply of Tap Water	5	1.108	1.395	0.728	0.529
Cultural, Educational and Sports Goods	6	1.075	1.718	1.545	2.312
Tobacco Processing	7	0.941	1.273	0.741	1.034
Electric Equipment and Machinery	8	3.050	3.548	4.331	4.120
Plastic Products	9	1.584	1.550	2.365	2.359
Printing and Record Medium Reproduction	10	0.601	0.687	0.957	1.190
Transport Equipment	11	6.536	10.97	4.872	5.331
Leather, Furs, Down, and Related Products	12	2.099	5.553	2.646	4.429
Furniture Manufacturing	13	0.994	1.653	1.010	1.366
Metal Products	14	2.537	5.329	3.043	3.268
Special-Purpose Equipment	15	3.266	2.942	3.153	2.875
Ordinary Machinery	16	3.861	3.633	5.140	3.837
Medical and Pharmaceutical Products	17	2.271	2.465	1.984	2.216
Logging and Transport of Timber and Bamboo	18	.	.	.	.
Food Processing	19	7.005	6.749	3.525	4.183
Timber Processing, Bamboo, Cane, Palm Fiber, and Straw Products	20	4.041	7.167	1.608	2.499
Food Manufacturing	21	2.356	3.306	1.929	2.244
Rubber Products	22	1.061	1.494	0.989	1.273
Textile Industry	23	6.923	6.198	7.477	7.263
Nonferrous Metals Mining and Dressing	24	1.660	1.927	1.756	3.576
Coal Mining and Dressing	25	22.86	27.65	13.27	16.69
Beverage Manufacturing	26	1.887	1.616	1.852	2.860
Nonmetal Minerals Mining and Dressing	27	1.293	1.505	0.960	1.240
Petroleum Processing and Coking	28	2.365	3.880	1.299	2.084
Smelting and Pressing of Ferrous Metals	29	4.957	10.44	6.054	9.755
Raw Chemical Materials and Chemical Products	30	5.002	3.677	6.642	5.035
Chemical Fiber	31	0.735	1.116	0.995	1.814
Production and Supply of Gas	32	0.381	0.335	0.317	0.346
Papermaking and Paper Products	33	1.776	1.708	1.892	1.531
Ferrous Metals Mining and Dressing	34	2.105	3.349	1.898	4.098
Smelting and Pressing of Nonferrous Metals	35	1.909	2.662	3.795	8.925
Nonmetal Mineral Products	36	6.406	5.072	7.624	6.347
Production and Supply of Electric Power, Steam, and Hot Water	37	4.962	4.492	4.223	3.454
Overall		3.536	7.393	3.250	5.423

policy.

3.4.2 Propensity-Score-Matching Model

Propensity-score matching can be used to solve the issue of endogenous treatment assignment. In this subsection, I introduce the propensity-score matching I used in this paper, which follows the model of Abadie and Imbens (2006). Let W_i indicates the treatment received by unit i: $W_i = 1$ if the city is designated as TCZ, otherwise $W_i = 0$. My main interest is the population average treatment effect:

$$\tau = E[Y_i(1) - Y_i(0)] \quad , \tag{3.1}$$

where Y is the variable of interest. In this case, Y will be industry pollution, economic activity variables, and industry composition measures. Under assumption (i), Y(0) is independent of treatment conditional on X, which is the observed characteristics, and under assumption (ii), support of X for the treated group (TCZ) is a subset of the support of X for the untreated group (non-TCZ), and the treatment effect for $X = x$ can be identified as:

$$\begin{aligned} \tau(x) &= E[Y(1) - Y(0)|X = x] \\ &= E[Y|W = 1, X = x] - E[Y|W = 0, X = x] \quad . \end{aligned} \tag{3.2}$$

The average treatment effect can be estimated by averaging $\tau(x)$ over the distribution of X:

$$\begin{aligned} \tau &= E[\tau(x)] \\ &= E[E[Y|W = 1, X = x] - E[Y|W = 0, X = x]] \quad . \end{aligned} \tag{3.3}$$

The distance of two observations in terms of covariates value can be calculated by using the standard Euclidean vector norm: $||x|| = (x'x)^{1/2}$. The mth closest unit to unit i is among those that have opposite treatment with unit i and indexed by $j_m(i)$. The X_j that is indexed with $j_m(i)$ would need to satisfy $W_j = 1 - W_i$, and

$$\sum_{l:W_l=1-W_i} I\{||X_l - X_i|| \leq ||X_j - X_i||\} = m \quad . \tag{3.4}$$

Let $J_M(i)$ denote the set of units for the first M matches for unit i: $J_M(i) = \{j_1(i), j_2(i), ..., j_M(i)\}$. Let $K_M(i)$ denote the number of times unit i is used as a match, given that we are matching units with the M closest units:$K_M(i) = \sum_{l=1}^{N} I\{i \in J_M(l)\}$. In this paper, I use the technology introduced in Abadie and Imbens (2006), matching with replacement. Abadie and Imbens (2006) show that matching with replacement produces better quality than matching without replacement by increasing the set of possible matches.

For each city, what I am interested in is $\tau_i = Y_i(1) - Y_i(0)$. However, for each city, I observe only one outcome, either with the treatment, $Y_i(1)$, or without the treatment, $Y_i(0)$. However, the missing outcome can be estimated using matching as

$$\hat{Y}_i(0) = \left\{ \begin{array}{cc} Y_i, & \text{if } W_i = 0, \\ \frac{1}{M}\sum_{j \in J_M(i)} Y_i, & \text{if } W_i = 1, \end{array} \right\} \tag{3.5}$$

and

$$\hat{Y}_i(1) = \left\{ \begin{array}{cc} \frac{1}{M}\sum_{j \in J_M(i)} Y_i, & \text{if } W_i = 0, \\ Y_i, & \text{if } W_i = 1. \end{array} \right\} \tag{3.6}$$

The estimator for the average treatment effect would be

$$\begin{aligned} \hat{\tau_M} &= \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i(1) - \hat{Y}_i(0)) \\ &= \frac{1}{N}\sum_{i=1}^{N}(2W_i - 1)(1 + \frac{K_M(i)}{M})Y_i \quad . \end{aligned} \tag{3.7}$$

3.4.3 Propensity-Score Calculation and Matching

To apply the propensity-score matching, I first need to calculate the propensity score. To do this, I need to estimate $Pr(TCZ = 1|X) = \Phi(X'\beta)$ using a probit model. I use cities' pollution emissions, area, population, GDP, GDP per capita, industrial output, and FDI in 1997 as X. Using X, I estimate cities' probability of being designated as TCZ in 1998—this is city's propensity score. I also set up a "caliper", to limit the maximum distance allowed. For each TCZ city, I'm matching it with the non-TCZ city with the closest propensity score within the caliper. I am using matching with replacement, which means that non-TCZ cities can get matched more than once. I use each city's number of matches as the weight. If the caliper is small enough, there will be cities without matches. For further analysis, I run a weighted regression excluding such cities without matches. I face a trade off between quality of matching and number of observations in the sample. As I tighten caliper,the quality of the match increased, which means cities in the control and treated groups are more balanced. However, more cities get unmatched with smaller caliper. I will discuss this further in the results.

3.4.4 Empirical Specification

Economic Activity

To study the causal effect of the TCZ policy on industrial SO_2 emissions, industrial SO_2 reduction, GDP growth, new FDI, population, and industrial output growth, I use the following specification:

$$y_{c,t} = \beta_0 + \beta_1 TCZ_c + \beta_2 X_{c,t} + \gamma_p + \delta_t + \epsilon_{c,t} \quad , \tag{3.8}$$

where $y_{c,t}$ is the variable of interest for city c at year t, TCZ_c is the indicator of whether the city is designated as TCZ or not, $\beta_2 X_{c,t}$ is the control variable for city c in year t, which includes the city's population, area, GDP, industry output, and FDI. I include γ_p to control for the province effect, and δ_t to control for the year effect. Since I am interested in how the Eleventh Five-Year Plan has affected the effect of the TCZ policy, I run this regression separately for the years before 2006 and for 2006 through 2010.

Industry Distribution

One possible unexpected consequence of the TCZ policy may be an effect on city-level industry distribution. Polluting industries may prefer to open in less regulated areas, while more regulated cities specialize in cleaner industries. I use firm-level data to construct the number of firms and employees in each industry sector for each city from 1998 to 2007. I analyze industry distribution in the following three steps:

- Step 1: I run the following specification using city-level data:

$$ln(y_{c,t}) = \beta_0 + \beta_1 TCZ_c + \beta_2 X_{c,t} + \gamma_p + \delta_t + \epsilon_{c,t} \quad , \tag{3.9}$$

  where $y_{c,t}$ is the number of firms or number of industry employees in city c in year t. From this specification, I can learn whether or not the TCZ is affecting the number of industry firms or industry employment.

- Step 2: I divide industries into clean sectors and dirty sectors based on their pollution intensity. For each city, I aggregate the number of firms and employment into dirty and clean sectors. I use the following empirical model:

$$\begin{aligned} ln(y_{p,c,t}) =& \beta_0 + \beta_1 TCZ_c + \beta_2 Polluting_p + \beta_3 TCZ_c \times Polluting_p \\ &+ \beta_4 X_{c,t} + \gamma_p + \delta_t + \epsilon_{p,c,t} \quad , \end{aligned} \tag{3.10}$$

  where $y_{p,c,t}$ is the number of firms or employment in each clean/dirty industry sector in city c in year t. $Polluting_p$ is the indicator for dirty sectors. β_1 measures whether there is a difference in number of firms or employment in clean industries, while $\beta_1 + \beta_3$

estimates those for dirty sectors.

- Step 3: Aggregating the number of firms and employment into city-sector levels for each year, I estimate the following specification:

$$\begin{aligned} ln(y_{i,c,t}) =& \beta_0 + \beta_1 TCZ_c + \beta_2 PollutionIntensity_i + \beta_3 TCZ_c \times PollutionIntensity_i \\ &+ \beta_4 X_{c,t} + \gamma_p + \delta_t + \epsilon_{i,c,t} \quad , \end{aligned} \tag{3.11}$$

where $y_{i,c,t}$ is the number of firms or employment for each industry sector i in city c in year t, and $PollutionIntensity_i$ is the pollution intensity for industry sector i. I use two measures for pollution intensity in this paper, one per value added and one per employee.

3.5 Results

In this section, I present the results of the estimation. First, I record the results for propensity-score as well as evidence of balancing after propensity-score matching. Next, I present results for the effect of the TCZ policy on pollution and economic activities. Then I include results on industry composition.

3.5.1 Propensity Scores

In my empirical strategy, I first estimate the propensity score for each city. The estimated result, using the probit model described in Section 3.4, is recorded in Table 3.6. Cities with higher industrial SO_2 emissions and higher GDP were more likely to get designated as TCZ. Controlled for other variables, including GDP and industry SO_2 emissions, more popular cities were less likely to be designated as TCZ. In Panel B of Table 3.6, I present the summary statistics for estimated propensity-scores for TCZ and non-TCZ cities. On average, cities in the TCZ have a higher propensity score.

Table 3.6: Propensity-Score Estimation Result (Probit Model)

Panel A: Estimation Results for Probit Model					
Variable		**Coefficient**		**Std. Err.**	
Area (1,000 km^2)		-17.374		26.674	
Industry SO_2 Emission (million tons)		0.522**		0.224	
Population (million)		-4.994**		2.215	
GDP (billion)		1.16**		0.514	
GDP/capita (1,000)		-0.149*		0.089	
Industry Output (billion)		-0.157		0.182	
FDI ($ million)		0.016		0.015	
Constant		1.495		0.956	
Province Fixed Effect		Yes			
N		141			
$\chi^2_{(29)}$		52.29			
Pseudo R^2		0.2817			
Panel B: Summary Statistics for Propensity-Score					
	N	Mean	Std. Dev.	Min.	Max.
Non-TCZ	52	0.427	0.220	0.000	0.921
TCZ	89	0.762	0.237	0.122	1

*: Significant at 90% level. **: Significant at 95% level. ***: Significant at 99% level.

Next, I match cities within the TCZ with non-TCZ cities based on their estimated propensity scores. I try out different values for the caliper, and test whether cities have balanced observed variations. To do this, I run a weighted regression of the variables of interest on the dummy of TCZ. The weight is calculated as

$$w_c = TCZ_c + (1 - TCZ_c)M_c \quad , \tag{3.12}$$

where TCZ_c is the indicator for TCZ cities and M_c is the number of matches for each non-TCZ city. Table 3.7 presents the mean for industrial SO_2 emissions, area, population, GDP, GDP per capita, industrial output, and FDI for TCZ and non-TCZ cities. This table also includes the p-value testing the null hypothesis that the mean is the same across TCZ and non-TCZ cities for each variable as well as for the joint test. When I shrink the caliper to 0.05 and to 0.01, almost all variables are balanced and the joint test cannot be rejected. For further empirical analysis, I will mainly set the caliper to 0.05.

Table 3.7: Balancing Check for Different Caliper Values

Caliper	0.3			0.1		
	Treated	Control	p-value	Treated	Control	p-value
Industry SO_2 Emission (million tons)	0.80788	0.37596	0.006	0.57921	0.44102	0.315
Area (1,000 km^2)	0.01047	0.01417	0.000	0.01048	0.01212	0.208
Population (million)	0.41911	0.488	0.039	0.33685	0.40274	0.073
GDP (billion)	3.6431	3.7034	0.862	2.2296	2.9593	0.011
GDP/capita (1,000)	8.5219	7.8914	0.319	7.0404	7.8811	0.274
Industry Output (billion)	6.0534	5.1784	0.165	3.3574	4.3977	0.035
FDI (million)	21.771	7.304	0.000	9.1099	7.1881	0.449
Off Support	23			50		
On Support	118			91		
p-value for Joint Test	0			0.207		

Caliper	0.05			0.01		
	Treated	Control	p-value	Treated	Control	p-value
Industry SO_2 Emission (million tons)	0.53834	0.48727	0.73	0.48444	0.44087	0.748
Area (1,000 km^2)	0.01088	0.01066	0.881	0.01061	0.01043	0.91
Population (million)	0.32449	0.34212	0.639	0.31462	0.32726	0.758
GDP (billion)	1.902	2.4302	0.05	1.8096	2.182	0.168
GDP/capita (1,000)	6.1592	7.8738	0.03	6.1183	7.5083	0.109
Industry Output (billion)	2.7816	3.8427	0.034	2.6926	3.5087	0.127
FDI (million)	5.2391	7.1057	0.133	4.6689	6.4061	0.188
Off Support	61			72		
On Support	80			69		
p-value for Joint Test	0.882			0.946		

3.5.2 Pollution and Economic Activities

The results for economic activities are recorded in Table 3.8 and Table 3.9. Each panel in these two tables consists of four columns. In the first column, the following specification is estimated:

$$y_{c,t} = \beta_0 + \beta_1 TCZ_c + \epsilon_{c,t} \quad . \tag{3.13}$$

Column 1 presents the OLS estimator with all the data, while Column 2 presents the propensity-score matching. Thus this specification is estimated using a weighted regression

Table 3.8: Main Results Table 1

	(1)	(2)	(3)	(4)
Panel A: Industry SO_2 Emission Change from Last Year				
TCZ	-857.450	505.821	4660.35**	-2653.62*
	(992.653)	(1273.473)	(2289.4)	(1459.22)
Years	2004–2012	2004–2012	2004–2005	2006–2010
R-squared	0.0004	0.1522	0.2183	0.0989
N	1928	770	192	385
PSM	N	Y	Y	Y
Panel B: ln(Aerosol Optical Depth)				
TCZ	0.0218*	-0.0529***	-0.0430*	-0.0726***
	(0.0122)	(0.0154)	(0.0222)	(0.0209)
Years	2000–2012	2000–2012	2000–2005	2006–2012
R-squared	0.6912	0.6703	0.6293	0.7160
N	2319	1114	578	536
PSM	N	Y	Y	Y
Panel C: Amount of SO_2 Emission Reduction				
TCZ	90.404***	86.990***	60.021***	106.657***
	(22.223)	(12.188)	(16.385)	(17.583)
Years	2004–2012	2004–2012	2004–2005	2006–2010
R-squared	0.0057	0.2617	0.1947	0.2988
N	1686	579	192	387
PSM	N	Y	Y	Y
Panel D: ln(GDP)				
TCZ	0.2122***	-0.0848***	-0.0460*	-0.0261
	(0.0195)	(0.0205)	(0.0245)	(0.0222)
R-squared	0.8518	0.8953	0.9020	0.9092
Years	2000–2012	2000–2012	2000–2005	2006–2012
N	2333	1114	578	536
PSM	N	Y	Y	Y

*: Significant at 90% level. **: Significant at 95% level.
***: Significant at 99% level.

with only the data on support. I use propensity score matching to estimate coefficients in all the following columns. I column 2, I add province and time fixed effects, as well as other control variables, including GDP increase, population increase, increase in the number of industrial firms, industrial output increase, and FDI. In the last two columns, I divide the data into before 2006 and 2006–2010. By doing this, I can study the effects of the TCZ play before and after the Eleventh Five-Year Plan. If I observe only the estimated coefficients from column 1 and column 2, the estimators in column 2 are always smaller in absolute value. This indicates that the TCZ policy is related to the error term, which also affects industrial emissions growth and economic activities. Therefore, without propensity-score matching, the estimator will be biased.

In Panel A, the dependent variable is the change in industrial SO_2 emissions. Compared with column 1, the estimated coefficient of the TCZ is much smaller in column 2, for which propensity-score matching is used. This is consistent with the conjecture TCZ cities may have a greater industrial SO_2 decrease because they are richer, more populated, and more developed, which may engender a higher demand for a clean environment. If I use the propensity-score matching, the estimate turns out to be positive, which indicates that, on average, matched TCZ cities have higher SO_2 emission during the whole period of analysis. There is no evidence that the TCZ policy is effective reducing industrial pollution in TCZ cities. However, if I consider only the 2000–2005 period the effect is positive and statistically significant. This means before the Eleventh Five-Year Plan, the TCZ policy, rather than driving cities to decrease industrial emissions, actually increased industrial emissions. The reason behind this is unclear. On the other hand, if I focus only on the Eleventh Five-Year Plan period, the effect is negative and significant—evidence that even though the TCZ may not have been very effective when it started, as the central government began to prioritize environmental protection, it became more effective in reducing industrial emissions.

In Panel B, the dependent variable is aerosol optical depth. The estimator is positive for a simple OLS estimation, which implies that TCZ cities are generally more polluted than non-TCZ cities. However, when using propensity-scores, I find statistically significant evidence that TCZ cities, compared with non-TCZ cities, have around 5.3% less aerosol optical depth for the 2000–2012 period. This effect is more pronounced when I consider only the period following the Eleventh Five-Year Plan(after 2006).

Another way to evaluate the effect of the TCZ policy on environment is by looking at the reduction of SO_2 emissions during the production process. I present these results in Panel C of Table 3.8. All results are positive and significant; TCZ cities reduce SO_2 emissions during the industrial production process more than non-TCZ cities. Moreover, if I compare the results in column 3 and column 4, the effect of the TCZ policy on SO_2 emissions reduction is much larger during the Eleventh Five-Year Plan than before.

GDP is the most representative measure of a city's economic growth. Therefore, by looking at GDP, I can observe the impact of the TCZ policy on economic growth. This result is recorded in Panel D of Table 3.8. If I consider the full 2000–2012 period (column 2), there is statistically significant evidence that TCZ cities have 8.5% lower GDP than non-TCZ cities. If I consider separately the periods before and after the Eleventh Five-Year Plan, the 2000–2005 period effect is 4.6%, the after-2006 period effect is smaller and not significant. However, the sign of the estimation stays the same over these three estimations. Thus, I conclude that TCZ cities sacrifice some GDP growth to achieve a better environment.

The effect of TCZ on cities' attractiveness in terms of FDI is recorded in panel E of Table 3.9. Overall, TCZ status makes cities less attractive to FDI. Overall, TCZ cities have 30% less new FDI than matched non-TCZ cities. This impressive effect may lead to a race to the bottom. Panel F and Panel G of Table 3.9 record effects of the TCZ policy on population growth and increase in industrial output, respectively. Over the 2000–2012 period, TCZ policy led to 4% less population in cities designated as TCZ. This effect is larger during the 2006–2012 period. Panel G shows no evidence that the TCZ policy significantly effected industrial output.

Overall, the TCZ policy's effect on economic activities is most significant during the Eleventh Five-Year Plan period. During this period, industry SO_2 emissions were smaller, ambient pollution levels were lower, and SO_2 emission reduction was larger in TCZ cities, while GDP growth and FDI were also smaller. Also during this period, TCZ cities attracted less population. This not only evidences that the TCZ policy is more effective when the central government cares more about environmental protection but also indicates that to achieve a better environment, cities have to sacrifice GDP growth and more stringent regu-

Table 3.9: Main Results Table 2

	(1)	(2)	(3)	(4)
	Panel E: ln(New FDI)			
TCZ	0.4022***	-0.2840***	-0.1479**	-0.1922***
	(0.0525)	(0.0554)	(0.0729)	(0.0530)
Years	2000–2012	2000–2012	2000–2005	2006–2012
R-squared	0.6822	0.7103	0.7798	0.7806
N	2315	1104	568	536
PSM	N	Y	Y	Y
	Panel F: ln(Population)			
TCZ	0.0032	-0.0412***	-0.0310**	-0.0555**
	(0.0122)	(0.0110)	(0.0156)	(0.0161)
Years	2000–2012	2000–2012	2000–2005	2006–2012
R-squared	0.8927	0.9341	0.9344	0.9314
N	2333	1114	578	536
PSM	N	Y	Y	Y
	Panel G: Increase in Industry Output			
TCZ	19.518***	2.646	-1.136	0.716
	(1.851)	(1.802)	(1.253)	(1.189)
Years	2004–2012	2004–2012	2004–2005	2006–2010
R-squared	0.0437	0.551	0.7923	0.7296
N	1822	783	193	392
PSM	N	Y	Y	Y

*: Significant at 90% level. **: Significant at 95% level.
***: Significant at 99% level.

lation makes cities less attractive to FDI and thus may also lose some of their labor force.

3.5.3 Industry Composition

I use city-level numbers of firms and industry-sector employment as measures for industry distribution. In table 3.10, I present industry-composition results for each step discussed in Section 3.4. Step 1 is the general study of the TCZ policy's effect on the number of industry firms and industry employment in the city. Step 2 and step 3 investigate the TCZ policy's effect on cities' industry-composition. In step 2, I divide the sectors into clean and dirty industries. In step 3, I use city-sector-level data.

Results for step 1 appear in Panel A of Table 3.10. The OLS estimates indicate that TCZ cities have a higher increase in the number of industry firms as well as in industry em-

Table 3.10: Industry Distribution

	Number of Firms		Employment	
		Step 1		
TCZ	0.1004**	-0.0714**	0.1339**	-0.0801**
	(0.0316)	(0.0316)	(0.0390)	(.0399)
Number of Observations	2025	2005	2011	1989
OLS	Y	N	Y	N
PSM	N	Y	N	Y
		Step 2		
TCZ	0.6685***	-0.1361***	0.6518***	-0.3709***
	(0.0402)	(0.0428)	(0.0487)	(0.0526)
Dirty Industry	0.6599***	0.4917***	0.7490***	0.4595***
	(0.0360)	(0.0324)	(0.0441)	(0.0404)
TCZ*Dirty Industry	-0.2823***	0.1288**	-0.1484***	0.5805***
	(0.0482)	(0.0520)	(0.0575)	(0.0617)
Number of Observations	4397	2005	4373	1989
Estimation Method	OLS	PSM	OLS	PSM
		Step 3		
TCZ	-0.0544***	-0.0273*	-0.01011	0.0423***
	(0.0165)	(0.0156)	(0.0218)	(0.0206)
Pollution Intensity	1.2***	0.0207***	3.345***	0.251***
($\times 10^{-4}$ for number of firms)	(0.15)	(0.0118)	(0.179)	(0.0142)
TCZ*Pollution Intensity	1.522***	0.123***	2.404***	0.156***
($\times 10^{-4}$ for number of firms)	(0.22)	(0.0177)	(0.257)	(0.0207)
Number of Observations	29759	29759	29386	29386
Value Added	Y	N	Y	N
Employment	N	Y	N	Y

*: Significant at 90% level. **: Significant at 95% level. ***: Significant at 99% level.

ployment. However, when I use propensity-score matching, the estimate becomes negative. This indicates that TCZ cities have fewer firms and less employment than non-TCZ cities. This is consistent with my finding on the economic activity results.

From the results of step 2, when I use the OLS estimates, TCZ cities have more firms and employment in clean industries. However, this difference become smaller for dirty industries. Propensity-score matching yields a totally different result. It turns out that TCZ cities have 13% more clean firms and 30% more employment in clean industries than non-TCZ cities. This difference is decreased dramatically for dirty firms, implying that instead of specializing in clean industries, TCZ cities have more dirty industries than non-TCZ cities.

In step 3, I use both measure of pollution intensity: emissions per value added and emissions per employee. The results are similar across different measures of pollution intensity, and the results for number of firms are similar to those for employment. The results from step 3 are consistent with those from step 2. I also observe that non-TCZ cities have more firms and greater employment in clean industries, while TCZ cities have more dirty industries than non-TCZ cities.

3.6 Conclusion

In this paper, I investigate the effects of the Two-Control-Zones policy on industrial SO_2 emission, pollution, economic activities, and industry composition. Further, I add the Eleventh Five-Year Plan to evaluate the change in the effects after the change in the government's attitude toward environmental protection. This paper uses propensity-score matching to solve the nonrandom TCZ assignment issue. My results show that since the Eleventh Five-Year Plan, the TCZ has been more efficient in terms of reducing industrial SO_2 emissions, and increasing the reduction in industry-generated SO_2. However, this change has come with a price. Cities in the TCZ have experienced lower GDP growth, less new FDI, and lower population. The effect of the regulation on industry composition is not exactly what we expect. It turns out that TCZ cities attract more firms in polluting sectors. This could be because polluting sectors are also capital-intensive and TCZ cities are capital-abundant. One limitation of this paper is that it doesn't provide implications about the mechanisms behind the economic activity change and especially industry composition. Further research could develop structural models to understand the counter-intuitive results regarding industry composition. Another limitation is that the data used in this paper for industrial emissions and economic activities cover only 2000–2012. To more accurately predict cities' chance of designated into TCZ, it would be helpful to observe data from before and after TCZ.

Bibliography

(2014, February). The effects of a minimum-wage increase on employment and family income. Technical report, Congressional Budget Office.

Abadie, A. and G. W. Imbens (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica 74*(1), 235–267.

Addison, J. T., M. L. Blackburn, and C. D. Cotti (2009). Do minimum wages raise employment? evidence from the u.s. retail-trade sector. *Labour Economics 16*, 397–408.

Allegretto, S., A. Dube, and M. Reich (2011). Do minimum wages really reduce teen employment? Accounting for heterogeneity and selectivity in state panel data. *Industrial Relations 50*(2), 205–240.

Allegretto, S., A. Dube, M. Reich, and B. Zippere (2017). Credible research designs for minimum wage studies: a response to neumark, salas, and wascher. *ILR Review 70*(3), 559–592.

Baskaya, Y. S. and Y. Rubinstein (2012). Using federal minimum wages to identify the impact of minimum wages on employment and earnings across the u.s. states. *Working Paper*.

Becker, R. and V. Henderson (2000). Effects of air quality regulations on polluting industries. *Journal of Political Economy 108*(2), 379–421.

Brochu, P. and D. A. Green (2013). The impact of minimum wages on labour market transitions. *The Economic Journal 123*, 1203–1235.

Burr, C. (2014). Subsidies, tariffs and investments in solar power market. *Working Paper*.

Card, D. and A. B. Krueger (1994). Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania. *The American Economic Review 84*(4), 772–793.

Chay, K. Y. and M. Greenstone (2003). Air quality, infant mortality, and the clean air act of 1970. *Unpublished*.

Chay, K. Y. and M. Greenstone (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*.

Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2013). The long-run impact of air pollution on life expectancy: evidence from china's huai river policy. *Unpublished*.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *39*(5), 829–844.

Cragoa, C. L. and I. Chernyakhovskiy (2017). Are policy incentives for solar power effective? Evidence from residential installations in the northeast. *Journal of Environmental Economics and Management 81*, 132–151.

Dong, C., R. Wiser, and V. Rai (2018). Incentive pass-through for residential solar systems in california. *Energy Economics 72*, 154 – 165.

Duan, N., J. P. Newhouse, C. N. Morris, and W. G. Manning (1983). A comparison of alternative models for the demand for medical care. *Journal of Business and Economic Statistics 1*, 115–126.

Dude, A., T. W. Lester, and M. Reich (2010). Minimum wage effects across state borders: estimates using contiguous counties. *Working Paper*.

Dude, A., S. Naidu, and M. Reich (2007). The economic effects of a citywide minimum wage. *ILR Review 60*(4), 522–543.

Foster, A. D., E. Gutierrez, and N. Kumar (2009). Voluntary compliance, pollution levels, and infant mortality in mexico. *American Economic Review 99*(2), 191–197.

Gillingham, K. and T. Tsvetanov (2016). Hurdles and steps: Estimating demand for solar photovoltaics. *Working Paper*.

Giuliano, L. (2013). Minimum wage effects on employment, substitution, and the teenage labor supply: evidence from personnel data. *Journal of Labor Economics 31*(1), 155–194.

Greenstone, M. (2004). Did the clean air act cause the remarkable decline in sulfur dioxide concentrations? *Journal of Environmental Economics and Management 47*, 585–611.

Greenstone, M. and R. Hanna (2014). Environmental regulations, air and water pollution, and infant mortality in india. *American Economic Review 104*(10), 3038–3072.

Henderson, J. V. (1996). Effects of air quality regulation. *The American Economic Review 86*(4), 789–813.

Hirsch, B. T., B. E. Kaufman, and T. Zelenska (2015). Minimum wage channels of adjustment. *Industrial Relations 54*(2), 199–239.

Hughes1, J. E. and M. Podolefsky (2015). Getting green with solar subsidies: evidence from the california solar initiative. *Journal of the Association of Environmental and Resource Economists 2*, 235–275.

Jardim, E., M. C. Long, R. Plotnick, E. van Inwegen, J. Vigdor, and H. Wething (2018). Minimum wage increases, wages, and low-wage employment: evidence from seattle. *NBER Working Paper*.

Meer, J. and J. West (2016). Effects if the minimum wage on employment dynamics. *Journal of Human Resources 51*(2), 500–522.

Neumark, D. and W. Wascher (2011). Does a higher minimum wage enhance the effectiveness of the earned income tax credit? *Industrial and Labor Relations Review 64*(4), 712–746.

Reich, M., K. Jacobs, A. Bernhardt, and I. Perry (2014, September). The mayor of los angeles' proposed city minimum wage policy: A prospective impact study. Technical report, Institute for Research on Labor and Employment University of California Berkeley.

Rogers, E. and S. Sexton (2014). Distributed decisions the efficiency of policy for rooftop solar adoption. *Working Paper*.

Sabia, J. J., R. V. Burkhauser, and B. Hansen (2012). Are the effects of minimum wage increases always small? New evidence from a case study of new york state. *ILR Review 65*(2), 350–376.

Sarzynski, A., J. Larrieu, and G. Shrimali (2012). The impact of state financial incentives on market deployment of solar technology. *Energy Policy 46*, 550–557.

Tanaka, S. (2015). Environmental regulations on air pollution in china and their impact on infant mortality. *Journal of Health Economics 42*, 90–103.

Thornberg, C., J. Levine, D. Schrader, B. Vanderplas, and M. Saia (2015, March). Cost-benefit analysis: Los angeles minimum wage proposal. Technical report, Beacon Economics.

Wicks-Lim, J. (2006). Mandated wage floors and the wage structure: new estimates of the ripple effects of minimum wage laws. *PERI working paper*.

Yang, S., Y. Chen, and G. M. Allenby (2003). Bayesian analysis of simultaneous demand and supply. *Quantitative Marketing and Economics 1*, 251–275.