The Dissertation Committee for Yaran Jin
certifies that this is the approved version of the following dissertation:

Essays on Regulations in the Electricity Industry

Committee:

Eugenio J. Miravete, Supervisor

Stephen P. Ryan

Robert Town

Thomas Wiseman

Haiqing Xu

Essays on Regulations in the Electricity Industry

by

Yaran Jin, Bachelor, M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2017

Essays on Regulations in the Electricity Industry

Publication No. ________________

Yaran Jin, Ph.D.
The University of Texas at Austin, 2017

Supervisor: Eugenio J. Miravete

In this dissertation, I study the welfare consequences of environmental and price regulations in the electricity industry. I address two important questions. One is that what are the welfare consequences of allowing firms to self-select between different types of environmental regulations. Another is how does the choice of transmission congestion pricing structure affect the emission externalities and fuel efficiency in the wholesale markets. I answer these questions by exploring several policy experiments in the state of Texas in the United States.

The first chapter is a general introduction to the Texas electricity industry and the conceptual framework of analysis in this dissertation. It consists of the institutional details of the industry, including market organizations, transmission congestion pricing structures, and emission regulations. Based on the institutional details, I discuss the theoretical implications and propose the empirical hypothesis for above research questions.

In the second chapter, I evaluate the welfare consequences of allowing firms to self-select between cap-and-trade regulation and intensity standards using the data from a unique voluntary NO_x emission cap-and-trade program in Texas from 2001 to 2005. The welfare evaluation focuses on the effects of such mixed policy instruments on emissions, industry profits, and market exit. I construct and estimate a structural model of power generating units equilibrium choices of policy instrument, emission abatement, and production to recover their abatement costs. With the estimated parameters, I simulate the equilibrium outcomes under a counterfactual mandatory cap-and-trade regulation. Results reveal that the mixed policy framework mainly benefits small and high-cost generating units. However, the aggregate emissions are lower and the aggregate profits are higher under the mandatory cap-and-trade regulation. I also document that the mixed policy instruments lead to a higher exit rate of older generating units.

In the year 2010, the Electric Reliability Council of Texas (ERCOT) changed from a zonal market structure to a nodal market structure to incorporate the cost of transmission congestion into the wholesale price. The third chapter compares the emission intensities and fuel efficiency of power generating units in the ERCOT before and after this regulatory change, to investigate its efficiency and environmental impacts in the congested areas. I find that the new nodal market structure has heterogeneous impacts on areas with different causes of transmission congestion. For counties located along the path to transferring wind generation from west Texas to east Texas, the nodal pric-

ing leads to increases in emissions from fossil-fired power plants, although the total increase in emission cost is not economically significant. Contrarily, the nodal pricing increases the fuel efficiency by 2-9.6% for power plants located around congested areas with excess load, and the estimated fuel cost saving is around $154.8m. The results provide important policy implications for future transmission network planning.

Table of Contents

List of Tables

List of Figures

Chapter 1

Markets Organizations and Regulations in the Texas Electricity Industry

The electricity industry provides the essential energy for economic development while subject to many kinds of economic and environmental regulations. The industry consists of four sectors: generation, transmission, distribution, and retail. Historically, in many countries, the industry was operated by regional vertically integrated natural monopolies and the government regulated the prices in every sector. Starting in the late 1990s and early 2000s, many states in the US restructured the industry and introduced competitive markets in the wholesale generation and retail sectors. For the restructured states, there were also many regulatory reforms about the market structure within the wholesale sector. On the other hand, the electricity generation consumes a huge amount of fossil fuels and emits many forms of pollutants to the environment, such as SO_2, NO_x, volatile chemical compounds and so on. These pollutants present big threats to human health but are also very costly to remove from the generation process. For this reason, regulating the emissions from this industry has important economic and environmental impacts.

As one of the most successful electricity market in the world, Texas has always been the focus of empirical research to improve our understanding of regulatory impacts in the electricity industry. This dissertation focuses on several regulatory changes in Texas in the 2000s to answer two research questions. One is that what are the welfare consequences of allowing firms to self-select between different types of environmental regulations. Another is how does the congestion price regulation affect the emission externalities and fuel efficiency of the wholesale markets. In this chapter, I introduce the institutional background of the Texas electricity industry and compare its feature with other states. At the end of this chapter, I lay out the conceptual framework of analyzing the above two research questions and provide the theoretical foundation for the empirical research in the following chapters.

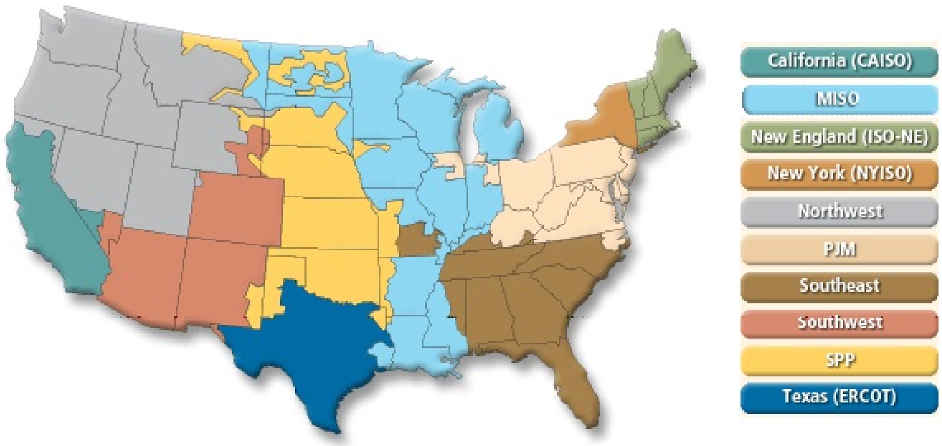
1.1 Texas Electricity Industry

1.1.1 Market Organizations

There are several regional electric power markets in the US (see figure 1.1). Some of them are restructured with well-organized competitive markets like those in northeastern states and Texas, but others are not. The electricity markets in Texas are the focus of this dissertation. For Texas, over 80% of the power generating units[1] belong to the Electric Reliability Council of Texas (ERCOT).

[1] A generating unit is a combination of fuel boiler and electricity generator. A power plant may own one or many generating units, and these units may have different fuel types, capacities and ages.

Figure 1.1: Electric Power Markets in the US



Source: FERC. https://www.ferc.gov/market-oversight/mkt-electric/overview.asp

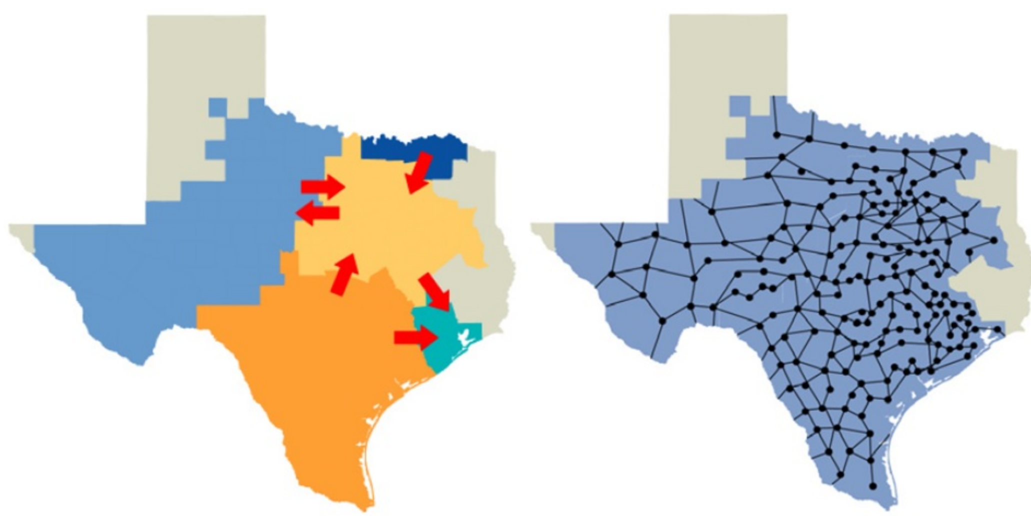
In the year 1999, the ERCOT became the first Independent System Operators (ISO) in the US. As of the year 2017, it manages 90 percent of the Texas state's electricity load and serves more than 24 million of Texas consumers. The ERCOT's function is to schedule the flow of electric power with more than 46,500 miles of transmission lines and 550 generating units[2]. It is one of the most competitive electricity markets in the North America. The wholesale generation market of the ERCOT was deregulated in the year 1995 and the retail market was deregulated in the year 2002. The total available generation capacity of ERCOT is more than 70,000 MW. The primary source of generation is the fossil-fuel source, especially natural gas. It also has a substantial amount of wind generation capacity in west Texas.

[2]Source: ERCOT website. http://www.ercot.com/about

Regarding economic regulations, this dissertation centers on the pricing structure of transmission congestion in the wholesale market. Transmission congestion is a critical issue in the electricity industry. It occurs due to the physical limits of the transmission lines and transformers. Whenever it happens, the electricity system is running at its full capacity and the lowest-cost generators cannot supply its electricity to the consumers. Such congestion is very costly and complicated to manage, and in each year the total congestion costs in the ERCOT are as large as hundreds of millions of dollars. Another contributing factor of congestion is that electricity energy is impossible to store without costly storage devices, which implies that the supply and demand have to equal to each other in the real-time market. To balance the market, the market regulator either uses potential inefficient administrative procedures to force supply to equal to demand from time to time, or incorporates the cost of congestion into market prices to let the participants adjust their strategies by themselves.

Before the year 2010, the ERCOT manages transmission congestion by a zonal market structure, and the electricity supply and demand is mainly balanced through pre-determined bilateral contracts between buyers and sellers. There is no day-ahead market but there is a real-time energy market, which covers less than 5% of total supply. The real-time energy market is divided into five loading zones: West, North, Northeast, South and Houston (see Figure 1.2). Each zone has its own market clearing prices for every 15-min interval. The suppliers and buyers are all charged with the same price in each zone. Given the zonal structure, there are two types of congestion: zonal congestions that happen when transmitting powers from one zone to another, and local congestions that happen within a single zone. The costs of zonal congestion will be added into the market clearing prices to provide signals to sellers and buyers in each zone to adjust their supply and demand, in order to alleviate the zonal congestions. However, when local congestion happens, the zonal price cannot provide any information about the geographic supply and demand to let sellers and buyers adjust their strategies voluntarily. In this case, the ERCOT cooperator will adopt complicated administrative procedures to guide power generating units to change their supply to balance the markets. The decision of how much supply provided by each generating unit might not be based on the cost-minimization of the whole system.

Figure 1.2: ERCOT Map



Source: ERCOT. Left: zonal map. Right: nodal map.

The Texas regulators established the zonal market structure because they believed that the main type of congestion in the ERCOT would be zonal congestions. However, the frequency of local congestions happened more than expected with tremendous costs. Consequently, in the December of 2010, the ERCOT started its nodal market structure to better manage the local congestion problem. The generators are paid with nodal prices (AKA locational marginal prices). Under the nodal market structure, the electricity grid in the ERCOT is further divided with more than 8,000 nodes, and each node's price will depend on the locational specific transmission capacity and nearby generators' marginal costs. I will define the nodal prices and discuss how to calculate them in the next subsection. However, the market is still divided into five loading zones and the consumers pay the zonal weighted average nodal price when purchasing the electricity in the wholesale market. The new market design also introduces a voluntary day-ahead forward energy market to help market participants schedule the supply and demand, as well as a congestion revenue rights market to help participants hedge against future congestion cost.

The regulatory change from a zonal to a nodal market structure is not limited to the ERCOT. For example, before April 2009, the California Independent System Operator (CAISO) adopted the zonal market structure. After that, the CAISO adopted the nodal market structure. A similar feature of the ERCOT and CAISO is that under the zonal market structure, these markets do not have a day-ahead market. After the implementation of nodal market structure, the day-ahead market starts operation in both markets. Currently, all the electric regional markets in the US with competitive wholesale markets adopts the nodal market structure.

As mentioned above, there are about 20% of electricity generation sources in Texas belong to other markets instead of the ERCOT. These markets include the Southwest Power Pool (SPP), Arizona-New Mexico-Southern Nevada (AZ-NM-SNV) subregion of the Western Electric Coordinating Council (WECC), and the Entergy subregion of the Southeastern Electric Reliability Council (SERC). The SPP didn't have organized real-time markets like in the ERCOT before 2007, and in March 2014 it also started operating a day-ahead market. Before 2014, the SPP mainly includes the states of Oklahoma, Kansas and the northern part of Texas outside of ERCOT. After 2014, it integrated the states of Nebraska, North Dakota and South Dakota. The Entergy subregion of the SERC didn't have the organized market structure before 2013. After 2013, this region was integrated into the Midcontinent Independent System Operator (MISO). Also, there are no organized markets in the AZ-NM-SNV subregion of the WECC. The electricity was mainly sold through bilateral contracts in this region.

1.1.2 A Introduction to Nodal Price

In this subsection, I will define transmission congestion and nodal and zonal prices with graphic examples. These examples are borrowed and modified based on the educational materials of the ERCOT[3], CAISO[4], the Independent Electricity System Operator in Canada[5], and Wolak (2014). These examples will be useful to generate empirical hypothesis for the following chapters.

1.1.2.1 Transmission Congestion

Several factors cause transmission congestion. One is the limited capacity of the transition line, and another is the law of electricity flows. Electricity flows through the transmission networks according to the Kirchhoff's circuit laws in physics. Instead of flowing according to the desires of supply and demand, electricity travels along the path of least resistance if there are multiple paths available in the transmission network.

Let me illustrate the Kirchhoff's circuit laws by considering an example of four locations with a seller in location A and a buyer in location B (see Figure 1.3 Case 1). All the four lines connecting the four locations have equal resistance. There are two paths for electricity to flows from A to B and the path A-D-C-B will have three times the resistance of path A-B. The flows from A to B will be divided into these two paths and the ratio of flow in each path is the inverse of their resistance. Suppose the seller A provides 100MW electricity to B without exceeding the transmission capacity, then 75MW will

[3]http://www.ercot.com/content/wcm/training_courses/14/nodal101_september_2013.pdf
[4]http://www.caiso.com/docs/2004/02/13/200402131607358643.pdf.
[5]http://docplayer.net/38818361-Nodal-pricing-basics-drew-phillips-market-evolution-program.html

flow on path A-B, and 25MW will flow on the other path. If there is another seller C, and suppose that A generates aMW and C generates cMW for buyer B(see Figure 1.3 Case 2). Then 75% of A's generation flows through A-B while 75% of C's generation flows through C-B. The net flows of each path also have changed and I summarize them in Table1.1.

Figure 1.3: Electricity Flows

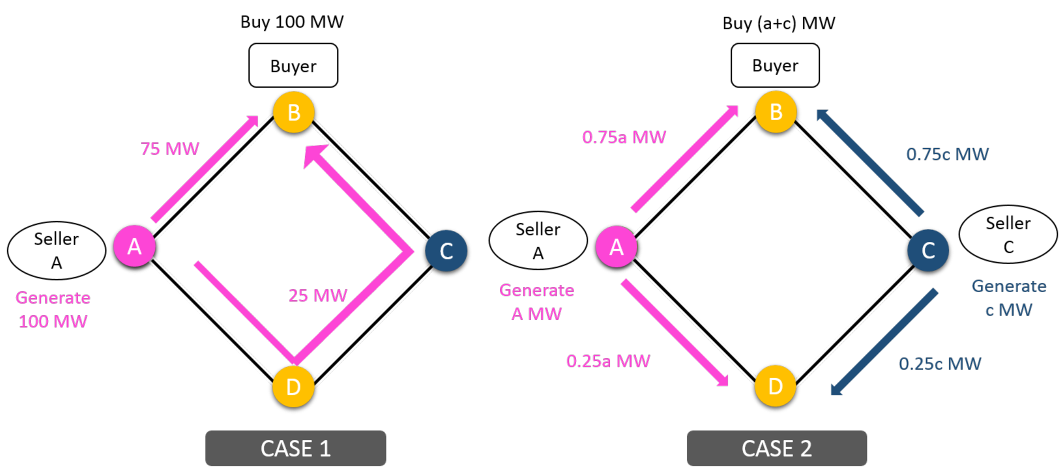


Table 1.1: Electricity Flows in Figure 1.3 Case 2

	Total Flow	From Node A	From Node C
A to B	0.75a+0.25c	0.75a	0.25c
A to D	0.25a-0.25c	0.25a	-0.25c
D to C	0.25a-0.25c	0.25a	-0.25c
C to B	0.25a+0.75c	0.25a	0.75c

The market operator dispatches different generation sources by the so-called merit order to minimize the total cost of electricity generation. The generators with the lowest cost will supply first until the flows along the transmission lines linking the lowest-cost generator reaching the capacity of the lines. In the above example, the total flows of each path in Table1.1 cannot exceed the capacity of each line. Whenever the total flows exceed the limited capacity, transmission congestion occurs. Then, the market operator has to deploy higher cost generator for supply to alter the electricity flows to satisfy the network constraint.

Transmission lines have both capacity constraints and transmission loss. Due to the existence of loss, the amount of electricity injected into one end of the line will not equal to the amount can be extracted at the other end of the line. In reality, the total electricity generation may have to exceed the total demand due to such loss. The main factor affecting the loss is the voltage of the line. The higher the voltage, the less the loss happens. This is why we always use high-voltage lines to transmit the electricity from wholesale sellers to the distribution firms. According to the EIA, in the US around of 6% of electricities are lost during the transmission process. However, to simplify the illustration, the examples in this chapter assume zero loss.

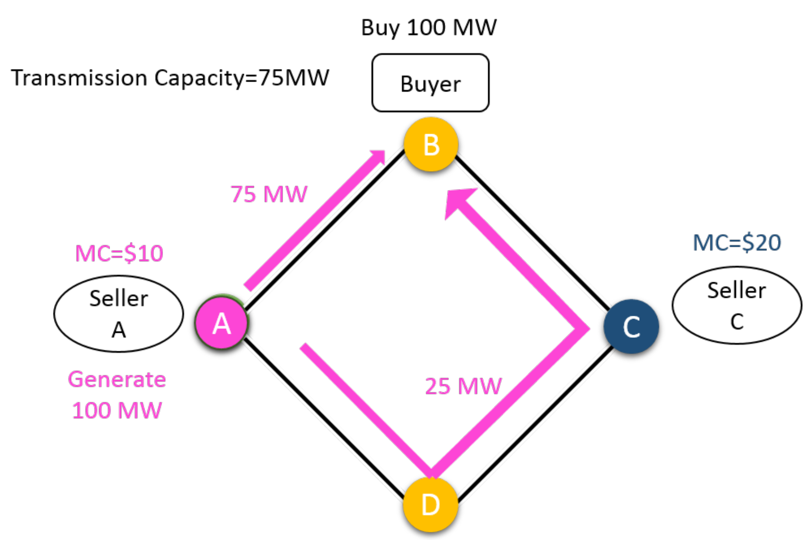
1.1.2.2 Nodal Prices with Congestion

This part illustrates how to calculate nodal and zonal prices by using the four-node example. Suppose the capacity for all the lines are 75MW in the examples, which means that the lines cannot transfer for any amount higher than the capacity. The dispatch order of the sellers will be determined by minimizing the total cost of generation. In the examples below, the generators

A and C have different marginal costs, and A's cost is lower than C (see Figure 1.4)[6]

The nodal price for each node is the marginal cost of supply at each node, which is the cost of supplying additional 1MW of electricity to a buyer at that node. Suppose the current market condition is that 100MW is already provided to the buyer at node B. The lower cost generator A should provide the amount as high as possible within the limitation of the transmission capacity. Under the current condition, A is actually able to supply all the electricity given the network constraint.

Figure 1.4: Transmission Constraint



[6]The wholesale markets are usually operated by the form of supplier auction. Let us assume that in the wholesale markets these sellers submit their bids representing their true marginal costs without any price manipulation. Also, I assume that there is no thermal loss of transmission in the following examples.

The nodal price of node A will still be the marginal cost of A, 10$, since A is able to provide the additional 1MW without causing any congestion. However, the calculation of nodal prices at B, C and D is more complicated. Let us start from finding the nodal price of node B. It is the same as considering the buyer at node B purchases 101MW. In this case, the seller at A cannot provide all the electricity, otherwise, it will violate the transmission constraint. We have to split the generation between seller A and C. Define the solution as A generating aMW and C generating cMW for the total 101MW. Then, (a,c) is the solution of the following linear programming problem to minimize the total cost of generation:

$$\min_{a,c}\ 10a + 20c\ s.t.$$

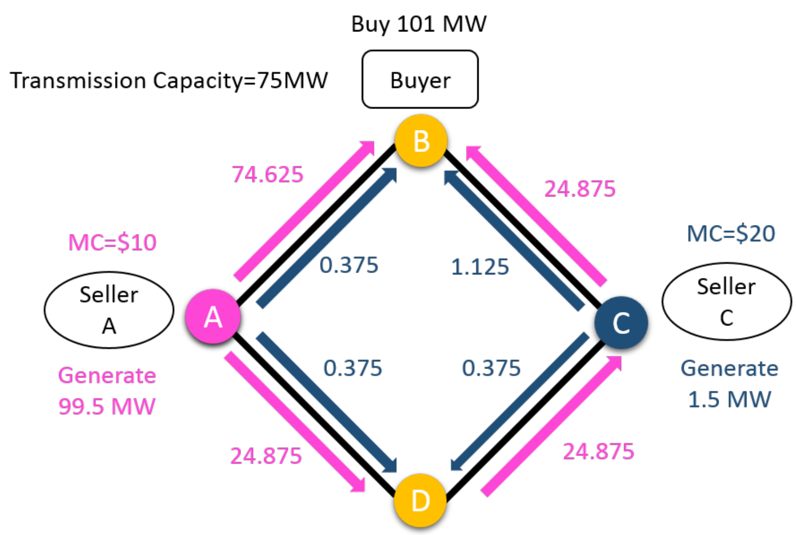
$$a + c = 101 \text{ (supply=demand)};$$

$$0.75a + 0.25c \leq 75 \text{ (the constraint of path A-B)}.$$

The solution is $(a, c) = (99.5, 1.5)$ (see Figure 1.5). Therefore, A provides 25% of the extra demand of 1MW at node B and C provides 75%. The nodal price at B is $(99.5 - 100) \times \$10 + \$1.5 \times 20 = \$25$.

To find the nodal price at node D, let us suppose there is a buyer at node D with 1MW of demand. Here the problem is complicated because we not only have to consider how power are split between A-B, A-D-C-B, C-B and C-D-A-B, but also how power are split between A-D, A-B-C-D, C-D and C-B-A-D. Define the solution as A generating aMW and C generating cMW

Figure 1.5: Nodal Price at B with Congestion



for the additional 1MW.Then, (a,c) is the solution of the following linear programming problem:

$$\min_{a,c}\ 10a + 20c\ s.t.$$

$$a + c = 1 \text{ (supply=demand)};$$

$$0.25a - 0.25c = 0 \text{ (constraint of path A-B)}.$$

The solution is $(a, c) = (0.5, 0.5)$ (see Figure 1.6) The nodal price at D is $(0.5) \times \$10 + \$0.5 \times 20 = \$15$.

To calculate the nodal price at node C, we can use the same procedure. Here A is unable to provide any additional electricity because the path A-B is already operated at full capacity. The additional 1MW demand at node C will be provided by seller C only. Therefore, the nodal price at C is 20$.

Figure 1.6: Nodal Price at D with Congestion

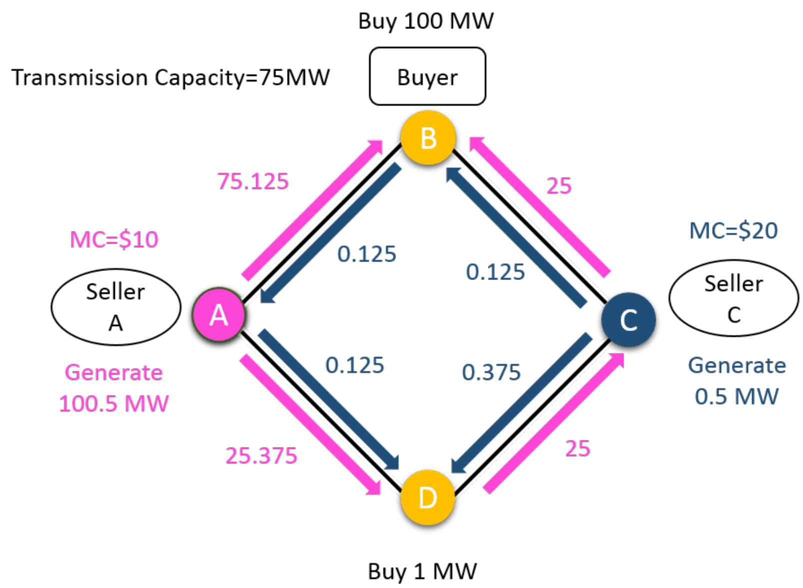
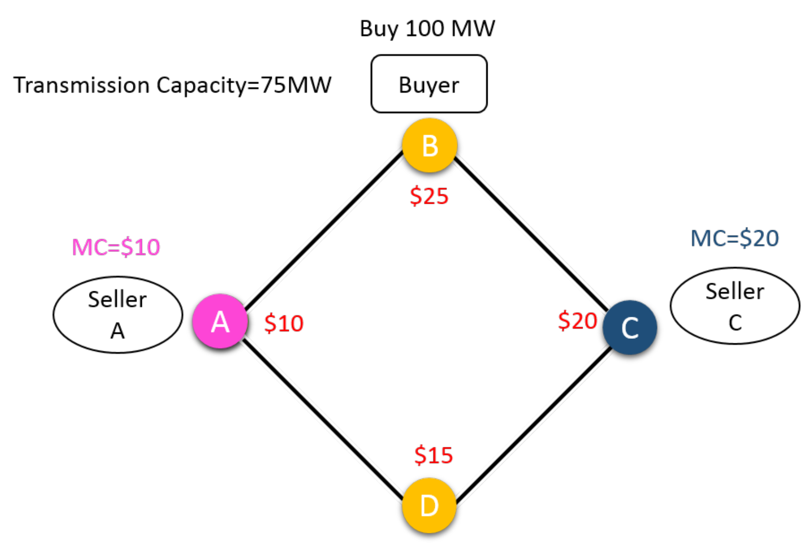


Figure 1.7: Summary: Nodal Prices with Congestion



To summarize, under the market condition that node B has 100MW demand, the nodal prices at each node is completely different (see Figure 1.7). The difference between the nodal prices at A and C is because of the difference in the marginal costs of the sellers. For node D, its nodal price is a weighted average of A and C because of the constraint of line AB and the split of the electric flows through the network. The nodal price at node B is the highest, which is because the line AB is already operated at full capacity and any additional demand would require low-cost seller A to reduce generation and high-cost seller C to increase generation.

1.1.2.3 Nodal Prices without Congestion and Zonal Prices

The above examples show how to calculate nodal prices with congestion. The calculation of nodal prices without congestion is very straightforward. Suppose the transmission capacity of the whole network is increased so that under current market condition any additional demand at any node will not cause any congestion on any line. Then whenever seller A is able to supply the additional demand, the nodal price will be the marginal cost of seller A. Otherwise, the nodal price will be the marginal cost of seller C.

Within each loading zone, the calculation of zonal prices will coincide with nodal prices without congestion. The zonal price is determined by the marginal cost of supply of the entire loading zone. Suppose all these four nodes belong to a single zone. If the marginal unit of supply is from seller A, then the zonal price is the marginal cost A. If the marginal unit of supply is from seller A, then the zonal price is the marginal cost C.

1.1.3 The Impacts of Congestion Pricing on Emission Externalities and Fuel Efficiency

In this subsection, I briefly discuss why congestion pricing market structure affects the emission externalities and fuel efficiency in the wholesale markets based on the four-node examples. As discussed previously, transmission congestion increases the cost of energy. Under the zonal market structure, the market is divided into separated zones and all the suppliers and consumers pay the zonal average marginal cost of production. Contrarily, under the nodal market structure, each supplier and consumer are represented by separated node with node-specific prices.

In theory, the nodal pricing is more efficient compared to the zonal pricing. The rationale behind is that the zonal pricing provides less incentive for generators to adopt cost-minimize behaviors. Under the zonal market structure, if congestion occurs the regulator will adopt administrative rules to decide which generator to provide electricity, instead of relying on the cost distribution of generators as in the nodal market structure. Moreover, the cost of production not only comes from fuel and labor costs, but also comes from the costs of reducing emissions under emission regulations. For example, in emission trading programs, the allowance prices present substantial additional costs to generators. Facing with additional cost imposed by emission trading programs, the nodal market structure gives generating units more clear signals of the demand and other suppliers' costs, and encourages generating units to use emission abatement technologies and fuel burning more efficiently. Particularly, if the market has many generation sources to compete in the wholesale markets, under the nodal market structure each individual will be more likely to improve their own efficiency to gain cost advantages against other competitors, so that the market operator would be more likely to dispatch them to minimize the total cost of the whole system. Therefore, we could propose the hypothesis that the nodal market structure reduces the emission intensities and improves the fuel burning efficiency for generating units.

However, if the transmission network is designed with flaws, the additional gains of the nodal market structure could not be realized. For example, if the most efficient and lowest-cost power plants are linking with low-capacity transmission lines, the network constraints will prevent them from supplying any additional quantity when the demand nearby is sufficiently high. In the previous four-node example, the limited capacity of the line AB and high demand cause the high-cost supplier C to replace the low-cost supplier A to generate electricity. In this case, the high-cost supplier C may not have the incentive to reduce marginal costs and improve efficiency. The construction of new lines should eliminate this problem, but the construction always involves outages of the lines during the construction process, adding further constraints into the network. In the empirical analysis in Chapter 3, I will show that how the network design and construction affects the performance of the nodal market structure.

1.2 Emission Regulations in the Texas Electricity Markets

The fossil-fueled power generating units are the major sources of various pollutants such as SO_2, NO_x, and CO_2. Different from SO_2 and CO_2 that directly come from the fuel content, NO_x mainly forms during the fuel combustion process. It contributes to ground-level ozone, atmospheric particles and acid rain that can cause serious damage to human health and the environment. Many policies have been implemented to reduce power plant NO_x emissions, and this dissertation focuses on the NO_x emission regulations in Texas during 2000 to 2014.

1.2.1 Regulations in the early 2000s

During 2000-2005, Texas adopted a mixture of cap-and-trade regulation and intensity standards regulation and a voluntary cap-and-trade provision to a subset of units. Figure 1.8 summarizes the regulatory framework.

1.2.1.1 Voluntary Cap-and-Trade Program

The cap-and-trade regulation targeted on the electricity generating units that were built before 1971. These units used to be exempted from the intensity regulations of the 1971 Clean Air Act and were allowed to operate without any emission abatement requirement. As a result, the emission intensities of these units were much higher than the post-1971 units. The Texas legislator was concerned that with the restructuring of electricity in-

Figure 1.8: Regulatory Framework

	West Texas	East Texas Intensity Standards
Pre-1971 Units	Mandatory Cap-and-Trade Cap=50% of 1997 emission	Mandatory Cap-and-Trade Cap=50% of 1997 emission
Post-1971 Units	Volunteer for Cap-and-Trade Cap=1997 emission	Volunteer for Cap-and-Trade Cap=intensity target* 1997 heat input
	Outside of Cap-and-Trade	Outside of Cap-and-Trade

Note: This figure summarizes the regulatory framework. All the units located in east Texas were subject to an intensity standard regulation. The pre-1971 generating units must participate in the cap-and-trade program to reduce half of their emissions relative to 1997. The post-1971 units could volunteer for this program, and the volunteer units' caps depend on whether they are subject to additional intensity standards regulation. For volunteer post-1971 units in the east, the baseline heat input to calculate emission cap is close to their 1997 heat input, but the specific choice of baseline heat input differs by county. For further details please refer to Appendix A.1. There are two emission markets in the cap-and-trade program: east Texas (in the light blue shade) and west Texas (in the gray shade).

dustry the power plant managers might shift more production to these older units, which would increase the total emissions from the industry. Therefore, in the year 1999, the Texas legislator established a cap-and-trade program to reduce 50% of the emissions from these pre-1971 units relative to their 1997 emission level.

The cap-and-trade program was mandatory for all pre-1971 units. There were two regional emission markets in this program (east and west Texas), and units were only allowed to trade emission allowances with units in the same region. The units had to ensure their annual emissions to be less than or equal to the allowances in their accounts within each control period, which was from the May of one year to the April of the following year. In each regional market, the allocated allowance on each pre-1971 unit was determined by this formula:

$$allocation = regional\ emission\ intensity * 1997\ heat\ input. \tag{1.1}$$

On average, each unit's allowance was around 50% of their 1997 emission. For the west region, the emission intensity baseline was set to be 0.195 lb/MMBtu, while for the east region the intensity baseline was 0.14 lb/MMBtu. The allocation for these units was fixed and remained constant ever since. This cap-and-trade program started in the May of 2003.

The cap-and-trade program was not limited to the pre-1971 units. The Texas regulator allowed the post-1971 units to volunteer for it. This volunteer provision was to give the generating units more flexibility to reduce emissions and lower total abatement costs. It was commonly acknowledged that the capital cost of retrofit and installing abatement devices on some of the pre-1971 units was very high. Accordingly, the regulator believed that when the volunteer units made extra emission reduction and sell the emission allowance to the pre-1971 units at a low price, the total cost of emission reduction in this program would decrease. The annual allowance allocation for the post-1971 units was either their 1997 emission or matched to other federal or state regulations on them. In the west Texas, because there was no extra emission requirement compared to 1997 on the newer units[7], their potential allowance in the cap-

[7]The previous paragraph mentions that the difference between pre-1971 and post-1971 units is that the post-1971 units subject to the entry restrictions in the 1970's Clean Air Act. However, based on the data, I find that the requirements of the Clean Air Act were very loose for the post-1971 units in the sample. The entrants after 2000 are exempted from all the regulations mentioned in this section. Therefore, how these requirements affecting the choice of entry is irrelevant and this paper will not explicitly model these requirements.

and-trade program was their 1997 emission. In the east Texas, the allowance for post-1971 was matched with the intensity standards regulation that will be discussed below. For the east post-1971 units, their emission allowance in the cap-and-trade program would equal to the product of their average heat input during 1996-1999[8] and the individual-specific intensity target. The volunteer decision had to be made before the September of 2000, which was before the start of the cap-and-trade program. The volunteer units could not withdraw from the program in the future once the decision was made. The post-1971 units decided not to participate in the cap-and-trade did not receive any emission allowances and could not buy and sell the allowances with other cap-and-trade participants.

In addition to requirements on NOx abatement, the cap-and-trade program also imposed SO_2 abatement targets on coal-fired pre-1971 units. However, there are only four coal-fired pre-1971 units in Texas. Most of the coal-fired units were built after 1971 and their SO_2 emission intensities were far below the threshold established by the Clean Air Act. Moreover, there were also very few volunteer coal-fired units. During the period 2000-2005, there were no other state regulations related to SO2 emissions except for this cap-and-trade program. Therefore, the analysis in the following chapters ignore this SO_2 requirement.

[8]The choice of baseline heat input differs by county. On average, the baseline heat input is close to their 1997 heat input. For further details please refer to Appendix A.1.

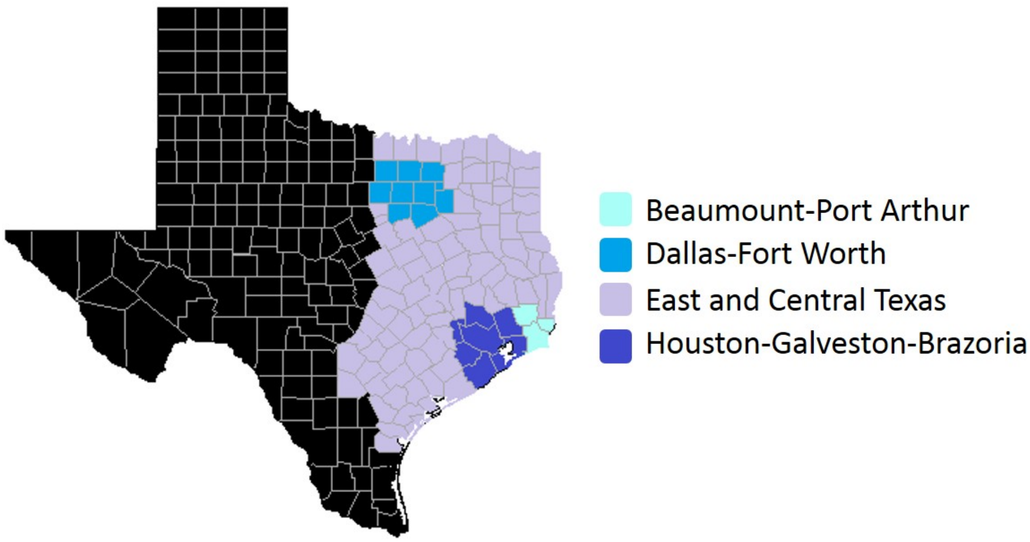
1.2.1.2 Intensity Standards in East Texas

In the east Texas, population and industrial activities are more concentrated compared to the west Texas. Many counties in east Texas violated or were very likely to violate the EPA ozone air quality standards in the 1990s. These counties were plotted in Figure 1.9. The counties located in the Dallas-Fort Worth, Houston-Galveston and Beaumont-Port Arthur areas had been classified as ozone non-attainment areas by the EPA. Because NOx emissions contribute to the formation of ozone, the Texas regulator under the requirement of the EPA issued a series of regulations, which were known as the State Implementation Plan, to control the industrial emissions generated from these counties. For the electricity generation industry, there was a regulation imposing county-specific thresholds on the NOx emission intensity. Table 1.2 divides the counties in east Texas into four regions and shows the maximum allowed intensities for each region. In some regions the intensity targets were also differentiated based on the fuel types. All the pre-1971 and post-1971 generating units in east Texas are located in these counties. For the post-1971 units, the last column in Table 1.2 shows the ratio of the cap-and-trade allowance over their 1997 emissions by region. If all the post-1971 units participated in the cap-and-trade program, on average they would have to achieve 44% of their 1997 emission level.

The enforcement of the intensity regulation in east Texas was different from the cap-and-trade program. If a regulated unit's emission intensity was higher than the target, it would face the penalty imposed by Texas Commission on Environmental Regulation regardless of its vintage. For the units subject to both the cap-and-trade program and the intensity regulation, they were considered as violating the intensity regulation if their emission was higher than the allowance in the cap-and-trade program and if their emission intensity was higher than the threshold in Table 1.2 at the same time. However, it did not imply that intensity standards exerted significant effects on every cap-and-trade units in the east Texas. The required compliance date for generating units are usually before the May of 2005, but for units located in the Dallas-Fort Worth region the required compliance date is after 2007. Because of this late compliance date for Dallas-Fort Worth, this paper assumes that during the sample period, the generating units in Dallas-Fort Worth did not include the violation penalty into their payoffs. Also, the intensity standards for the region East and Central Texas are almost the same as the intensity baseline used in the cap-and-trade program for post-1971 units. Therefore, the overlapping regulations mainly affect units located in the Beaumont-Port Arthur and Houston-Galveston areas.

The penalty on intensity violation consisted of two parts. The first part was the penalty on avoided fixed costs. The violating units usually failed to install an effective abatement technology to avoid high fixed costs. Upon the violation, these units had to pay half of the difference between the fixed cost of the technology they had installed and the fixed cost of a more effective technology, which could be as large as several million dollars. The other portion of the penalty depended on the degree of violation and the length of

Figure 1.9: East Texas Counties with Power Generating Units



Note: This figure depicts the counties in east Texas with power generating units in the sample. County-specific intensity standards are specified in Table 1.2.

Table 1.2: Intensity Standards for East Texas Power Generating Units

Region	Intensity Target (lb/MMBtu)	Emission Cap/1997 emission
Dallas-Fort Worth	0.03	0.51
Beaumont-Port Arthur	0.1	0.80
Houston-Galveston	Gas 0.03; Coal 0.05	0.23
East and Central Texas	Gas 0.14; Coal 0.165	0.48
All Planning Counties		0.44

Note: This table shows the intensity targets of each region that were published in *Texas Register* during 1999-2001 before the cap-and-trade volunteering decision was made. *Texas Register* also listed the rule to convert the intensity target into emission quantity target for generating units assuming these units maintained the average heat input level during 1996-1999. The last column is calculated as the ratio of the sum of the emission quantity targets over the sum of the 1997 emissions of all the post-1971 units in each region. Not all the post-1971 units in east Texas participated in the cap-and-trade program (see next section) so the last column is only to show what their emission allowances would be compared to the base year hypothetically, but does not represent their actual emissions.

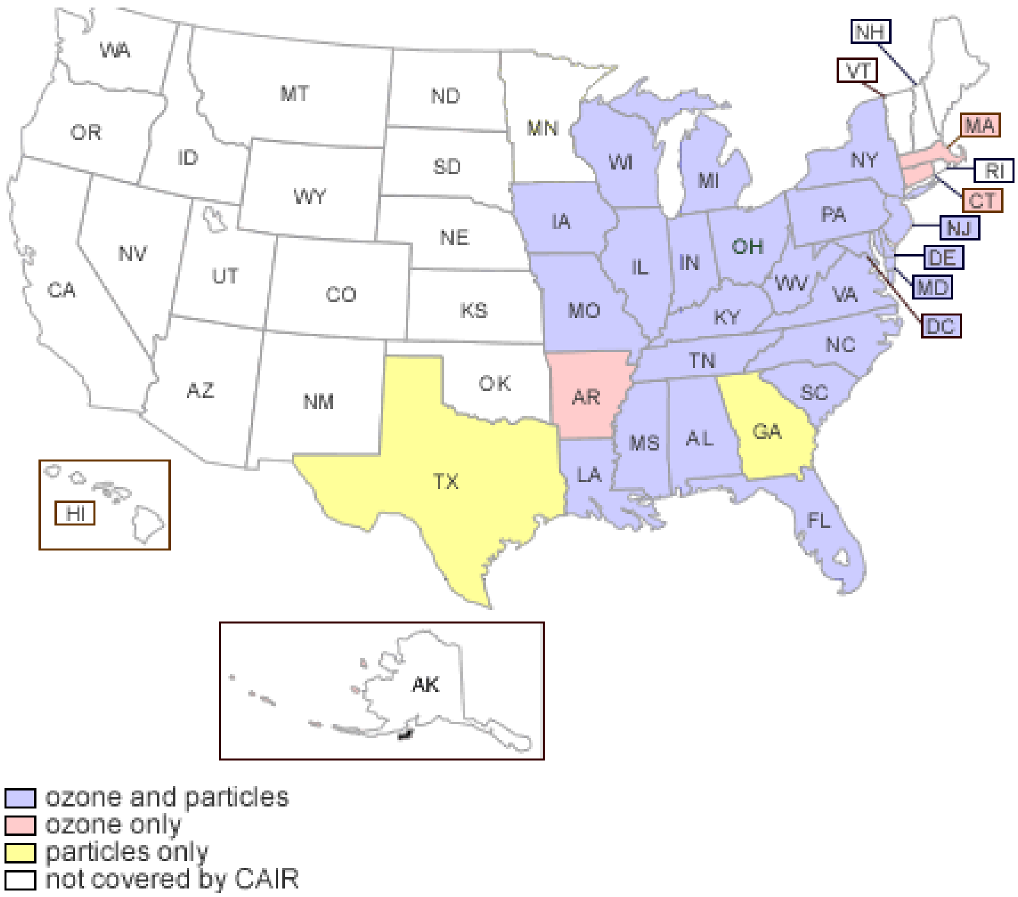
violation, which ranged from \$0 to \$25,000 per violation per day as stated in the penalty policy rule. However, because the enforcement office usually imposed penalties on air quality violation together with other violations such as the violation of water quality regulation, the actual penalty data on the NOx emission regulations was unavailable. According to the penalty rule that this variable penalty tax is proportional to the degree of violation, and that the length of violation is proportional to the quantity of production, this paper will propose a parameter representing the average value of penalty tax rate faced by violating units in the model and estimate its value from the data.

There were also other exemptions from this regulation for generating units satisfying certain criteria. These exemption rules are taken from *Texas Register* published on March 12th in 1999, May 5th in 2000 and January 12th in 2001. In regions except for Dallas-Fort Worth, units with annual heat input less than 22,000 MMBtu are exempted. In the regions of Houston-Galveston and Beaumont-Port Arthur, units start commercial service after 1992 are exempted. In the east and central Texas, units start commercial service after 1995 are exempted. Moreover, based on the rules of regulations, if a unit does not participate in the cap-and-trade program, and if the quantity of its emission is lower than a quantity emission target, it also can avoid paying the penalty. For units located in Houston-Galveston, their quantity target is the product of the intensity target and their average heat input during 1997-1999. For other units in the east Texas, the quantity target is the product of the intensity target and their average heat input during 1996-1998. The required compliance year for generating units are usually before the May of 2005, but for Dallas-Fort Worth the required compliance date is after 2007. This paper assumes that during the sample period, the generating units in Dallas-Fort Worth did not account for the violation penalty payment into consideration because of this late compliance date. Therefore in this paper, none of the generating units in Dallas-Fort Worth are treated as violating the intensity standards.

1.2.2 Regulations during 2006-2014

NO_x and SO_2 could travel long distance with the wind so that emissions generated from one state will contribute to the emission damage to human health in a faraway state. To deal with this problem, EPAs Clean Air Interstate Rule required 28 eastern states to make reductions in sulfur dioxide (SO_2) and nitrogen oxides (NO_x) emissions through emission trading programs. The legislation was passed in the year 2006 and was in action from 2009 to 2014. It covered 28 states in the east of the US (see Figure 1.10). It has two categories: one is for yearly emission and another is for the ozone season. Texas is regulated by the yearly program only. It was estimated by the EPA that the CAIR helped reduce NO_x and SO_2 emissions from Texas by 25% and 39% respectively. The NO_x emission trading started from 2009 and replaced the State Implementation Plan for the electricity industry in its covering states. The SO_2 emission trading started from 2010 and replaced the Acid Rain program in its covering states. After the year 2015, EPA changed the requirements of the CAIR and implemented new programs under the name of Cross-State Air Pollution.

Figure 1.10: CAIR Map



Source: EPA

All fossil fuel generating units in Texas are required to participate in the CAIR and they are divided into two types to receive the allocations of emission allowances. For those units starting commercial operation before 2001, their allowance is given by a complicated formula accounting for their average heat input in 2001-2004 and their fuel type (*30 TAC Chapter 101, Subchapter H, Division 7*). For the other type of units starting operation after 2001, their allowance is given by the Texas Commission on Environmental Quality case by case.

1.2.3 The Welfare Consequences of Mixed Policy Instruments with Self-Selection

This section provides conceptual analysis and graphic examples to illustrate why the mixed policy framework in Texas during 2000-2005 could lead to different outcomes compared to a mandatory cap-and-trade policy. I will discuss the trade-offs facing the generating units under different policy instruments and how the cost-effectiveness of the mixed policy framework relies on the distribution of abatement costs and targets. Without going into the details of policy backgrounds and mathematical modeling, this section emphasizes the fundamental features of the policy instruments and provides the framework for empirical analysis.

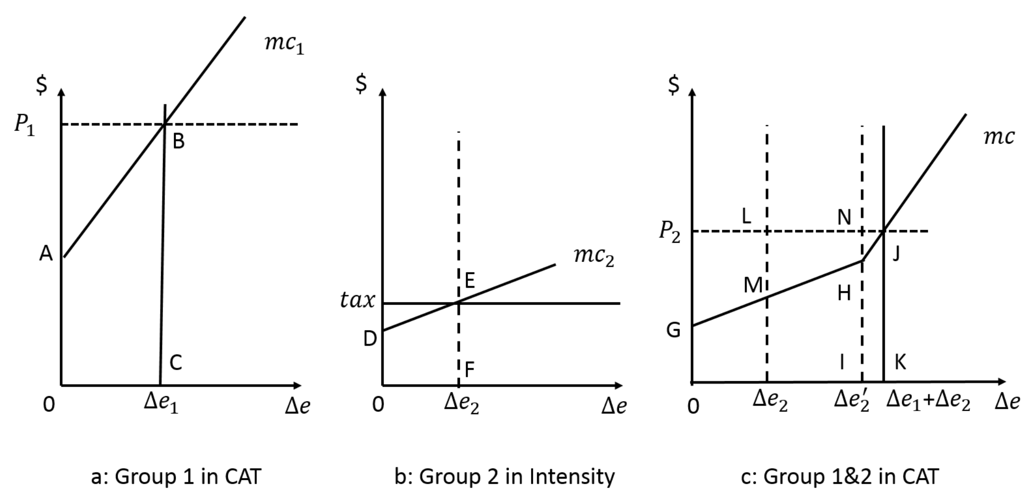
First of all, because the cap-and-trade regulation and the intensity standard regulation offers individual units with different incentives, the mixture of policy instruments can generate different outcomes compared with a uniform instrument. In the cap-and-trade program, units have to hold enough emission allowances to cover their quantity of emissions. They can achieve emission abatement by production reduction, investment in abatement technology or emission market transaction. Assuming that the units in the cap-and-trade market have different marginal costs of emission abatement, some low-cost units could reduce their emissions to a level below its allocated allowances and sell the extra allowances to other high-cost units at a low price. Then, the total abatement costs of the cap-and-trade market could be lowered by such interactions between the participating units. Contrarily, under the intensity standards regulation there is no emission market where units can interact with each other. Units have to make the trade-off between investing in abatement technology to limit their emission intensity or paying penalty taxes to the regulator upon violation. Given this difference between the cap-and-trade and intensity standards, those units with the voluntary cap-and-trade provision might make different choices under these two regulations. Their participation decisions will further affect the abatement choices by the mandatory cap-and-trade participants. Therefore, the outcomes of voluntary cap-and-trade programs depend on which set of units participate in the cap-and-trade market.

The most distinct difference between the voluntary cap-and-trade programs and a mandatory cap-and-trade policy is that the total emissions under the former are uncertain before the participation decisions are made, whereas the total emissions are fixed by the total allowance allocation under the latter. The voluntary cap-and-trade programs are targeting on the mandatory participants to reduce their quantity of emissions, implying that only the summation of the mandatory participants' emissions is fixed. For the rest of the units, if they decide to volunteer for the cap-and-trade, their emissions are limited by the allocated allowances, but if they decide to choose the intensity regulation instead, the intensity regulation does not bind the absolute quantity of emissions. In the equilibrium, if some units decide to volunteer while others do not, the total emissions and cost efficacy of the voluntary cap-and-trade program would be completely different from a mandatory cap-and-trade policy.

The first example is to show adding the voluntary provision could achieve the same optimal outcome as a mandatory cap-and-trade regulation if the average abatement cost of potential volunteers is significantly lower than the mandatory participants. In this example, the units are divided into two groups. The first group is mandated to participate in the cap-and-trade program to cut their emissions, and has the marginal abatement cost curve mc_1 as shown in Figure 1.11(a). These mandatory participants are given abatement target Δe_1 by the regulator. Hence, without any volunteers, the equilibrium emission price in the cap-and-trade market will be P_1 as shown in Figure 1.11(a). The second group of units has the marginal abatement cost curve mc_2. This group could choose to volunteer for the cap-and-trade program, and if they do not volunteer, they are subject to the partial tax regulation as shown in Figure 1.11(b). The abatement target for the second group is Δe_2. An implicit assumption in Figure 1.11(b) is that the tax rate equals to the second group's marginal cost of abatement at Δe_2, so that under the tax regulation the second group's abatement is the same as the target imposed by the regulator. Figure 1.11(c) shows the outcomes when both groups are in the cap-and-trade market. Here the aggregate marginal abatement cost curve is the horizontal summation of mc_1 and mc_2, and the equilibrium emission price

P_2 is determined by the intersection of the marginal cost curve and the sum of the abatement targets ($\Delta e_1 + \Delta e_2$). At P_2, the second group will abate $\Delta e_2'$, and it will sell the extra allowances ($\Delta e_2' - \Delta e_2$) to the mandatory group. Since the second group earns positive profits (the area of region LMNH) of selling allowances in the cap-and-trade market, it is optimal to volunteer for the cap-and-trade. Therefore, the total abatement costs decrease compared to the case without the voluntary provision, and the first group also benefits by avoiding the high cost of abatement. Table **??** shows the detailed welfare decomposition of this example. Here adding the voluntary provision to the second group leads to the same results as the mandatory cap-and-trade scenario.

Figure 1.11: Graphic Example 1: Benefits of the Voluntary Provision



Note: (a) shows the abatement choice of Group 1 when only itself participates the cap-and-trade market. (b) shows the abatement choices of Group 2 under the intensity tax regulation. (c) shows the result if both groups are in the cap-and-trade market.

Table 1.3: Welfare Decomposition of Example 1 and Figure 1.11

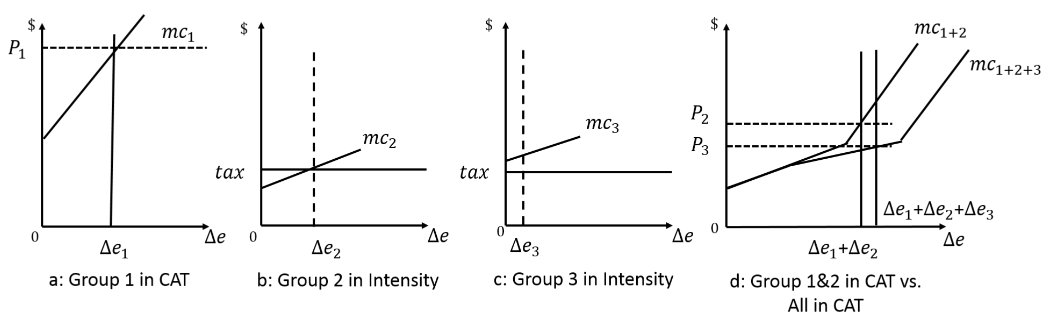
	Abatement	Abatement cost	Individual Net cost
Panel a: No Voluntary Provision to Group 2			
	Abatement	Abatement cost	
Group 1	Δe_1	OABC	
Group 2	Δe_2	ODEF	
Total	$\Delta e_1 + \Delta e_2$	OABC+ODEF	
Panel b: With Voluntary Provision to Group 2			
	Abatement	Abatement cost	Individual Net cost
Group 1	$\Delta e_1 + \Delta e_2 - \Delta e_2'$	HIJK	HIJK+$P_1(\Delta e_2 - \Delta e_2')$
Group 2	$\Delta e_2'$	OGHI	OGHI-$P_1(\Delta e_2 - \Delta e_2')$
Total	$\Delta e_1 + \Delta e_2$	OGHI+HIJK	OGHI+HIJK

Note: This table shows the amount of emission abatement and costs by each group of units in Example 1. The column 'abatement cost' is the integral of the marginal abatement costs from zero to the amount of abatement made by the units in each row. The column 'Individual Net Cost' represents the sum of the abatement costs and the cost of buying emission allowances in the cap-and-trade market for each individual group. The Panel b's result is the same as the mandatory cap-and-trade as the group 2 decides to volunteer.

If there is another group of units with a different abatement cost curve, adding the voluntary cap-and-trade option might yield different results compared with the mandatory scenario. Let this new group be group 3 with the marginal cost curve mc_3 as shown in Figure 1.12(c). This group is originally subject to the partial tax regulation and given the option to volunteer for the cap-and-trade program. The assumptions are that the abatement target imposed by the regulator on this group is Δe_3 and that the tax rate is the same as for group 2. Based on Example 1 it is trivial to show that group 2 will volunteer for the cap-and-trade market regardless of group 3's decision. As the tax rate is always below group 3's marginal cost, under the intensity tax regulation this group will not make any abatement and will pay the regulator ($tax\ \Delta e_3$) for its violation. However, when this group is included in the cap-and-trade market, because mc_3 is lower than mc_1, and because the total of the abatement targets ($\Delta e_1 + \Delta e_2 + \Delta e_3$) is higher than the maximum of abatement could be achieved by group 2, group 3 will no longer make zero abatement. In the mandatory cap-and-trade scenario, the equilibrium price will be P_3 as shown in Figure 1.12(d). Figure 1.13(a) shows that group 3 will abate $\Delta e_3^{'}$ and sells the extra allowances ($\Delta e_3' - \Delta e_3$) at P_3. Its net cost in the cap-and-trade will be the area of region OQVW plus the allowance revenue $P_3(\Delta e_3^{'} - \Delta e_3)$. Comparing with its net cost of ($tax\ \Delta e_3$) under the tax regulation, it could lose extra money participating in the cap-and-trade market. If the area of the region QYSW is larger than the area of the region XVS, it will not volunteer for the cap-and-trade market. Meanwhile, the aggregate social welfare of the mandatory cap-and-trade scenario might be larger than the voluntary scenario. As having explained in Figure 1.13 and Table **??**, in the mandatory scenario, the total amount of abatement is increased by Δe_3 and the aggregate abatement costs could also be reduced by the amount represented by the difference between the areas of the region RHIZ and the region ZYPK. Assuming the benefit of abatement is increasing in the quantity of abatement and the tax penalty can be freely transferred to people suffering from emission damage, it is very likely that the aggregate of welfare in the mandatory participation scenario is higher. The reason is that when group 3 makes the choice of participation, it does not include its positive externality of emission abatement and cost saving in the mandatory scenario into its own decision process. Such conflicts between individual and social welfare imply that the voluntary provision to group 3 might yield inefficient aggregate outcomes.

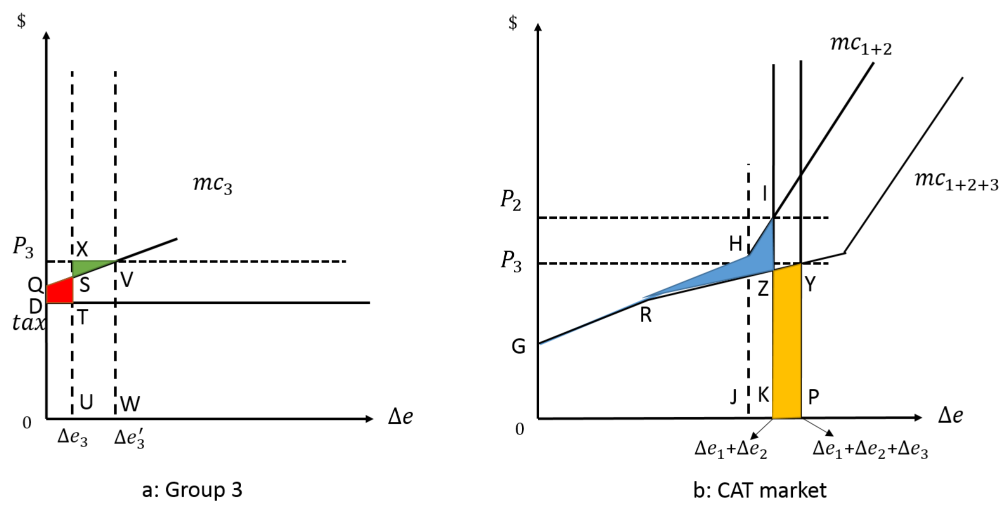
The examples in this section illustrate that the mixed policy framework could benefit a subset of generating units at the expense of higher total emissions and higher total abatement costs. The fundamental cause is that the individual cap-and-trade participation choice does not fully internalize all the abatement benefits in the mandatory regime. As a result, the aggregate welfare would not align perfectly with individual welfare. The problem is complicated by the dependence of the participation decision on the distributions of individual abatement costs and targets. The next chapter will use a structural approach to empirically examine the welfare consequences of the mixed policy framework.

Figure 1.12: Graphic Example 2: Inefficiency of the Voluntary Provision



Note: (a) shows the abatement choice of Group 1 when only itself participates the cap-and-trade market. (b) shows the abatement choices of Group 2 under the intensity tax regulation. (c) shows Group 3 makes no abatement under the intensity tax regulation despite the abatement target Δe_3. (d) compares the results in the cap-and-trade market with and without Group 3.

Figure 1.13: Welfare Decomposition in Example 2



Note: (a) shows the abatement choices of Group 3 in the cap-and-trade market and compares its cost with the intensity regulation. (b) shows the aggregate emission and costs in the cap-and-trade markets. The marginal cost curve on the left is the horizontal summation of the marginal costs of Group 1 and Group 2, while the marginal cost curve on the right is the horizontal summation of the marginal costs of all the groups.

Table 1.4: Welfare Decomposition of Example 2 and Figure 1.13

Panel a: Voluntary Cap and Trade Provision for Group 2,3			
	Abatement	Abatement cost	Individual Net cost
Group 1&2	$\Delta e_1 + \Delta e_2$	OGHIK	OGHIK
Group 3	0	0	tax$*\Delta e_3$
Total	$\Delta e_1 + \Delta e_2$	OGHIK	
Panel b: Mandatory Cap and Trade for all Groups			
	Abatement	Abatement cost	Individual Net cost
Group 1&2	$\Delta e_1 + \Delta e_2 +$ $\Delta e_3 - \Delta e_3'$	OGRYP-OQVW	OGRYP-OQVW $+P_3(\Delta e_3' - \Delta e_3)$
Group 3	$\Delta e_3'$	OQVW	OQVW-$P_3(\Delta e_3' - \Delta e_3)$
Total	$\Delta e_1 + \Delta e_2 + \Delta e_3$	OGRYP	

Note: This table shows the amount of emission abatement and costs for each group of units in Example 2. Based on Example 1 it is trivial to show that the group 2 will volunteer for the cap-and-trade regardless of the decision of the group 3. Therefore, the panel a's result is based on the participation decision by the group 2. The column 'abatement cost' is the integral of the marginal abatement costs from zero to the amount of abatement made by the units. The column 'Individual Net Cost' represents the sum of the abatement costs, the cost of buying emission allowances in the cap-and-trade market, and the individual payment for intensity violation if not in the cap-and-trade for each individual group. The total net cost is not shown in this table because under the assumption that the intensity tax penalty can be transferred to the public victims of emissions at no additional cost, the total net cost should be the same as the total abatement cost.

Chapter 2

The Welfare Consequences of Mixed Policy Instruments within Voluntary Emission Cap-and-Trade Programs

2.1 Introduction

This chapter studies the choices of NO_x emission control policy instrument in the electricity generation industry by exploring a unique regulatory framework in Texas in the early 2000s. Different from the conventional approaches to regulating every power generating units by a mandatory regional cap-and-trade regulation or technology mandate, the Texas regulator adopted a mixture of the cap-and-trade regulation and intensity standards selected by power generating units.

During the period from 2000 to 2005 in Texas, the power generating units built before 1971 were mandated to enroll in a cap-and-trade program to reduce emissions. Other generating units built after 1971, which were originally subject to federal or state-level intensity standards, could volunteer for this cap-and-trade program. At the same time in the east Texas, there is a separate intensity standards regulation imposed on generating units regardless of their vintage, and this regulation partially overlaps with the cap-and-trade regulation. Under such a vintage and spatially differentiated regulatory framework, there was actually very little voluntary participation into the cap-and-trade program by the post-1971 units. Meanwhile, a large proportion of pre-1971 units has exited from the market since the year 2000. Moreover, generating units regulated by different policy instrument made distinct choices of abatement technology investment and production. These observations motivate this chapter to propose the research questions that when compared with a mandatory cap-and-trade regulation, what would be the gains or losses in emissions and industry profits by adopting such a voluntary framework with mixed policy instruments, and whether the mixed policy instruments could accelerate or retard the closure of pre-1971 units.

The research question and results of this chapter can draw important policy implications. Despite the complexity of the Texas regulatory framework, it includes a range of possible emission regulations at broader regional levels like the U.S. or the European Union. Taking the U.S. as an example, in recent years the EPA issues a series of federal emission regulations on the electricity generation industry, such as the Clean Air Interstate Rule with a mandatory cap-and-trade instrument and the Clean Power Plan with a mixture of state-specific instruments. Sometimes the introduction of mixed or differentiated policy instruments may not stem from the cost-benefit perspective, but due to particular juridical or political considerations. There are a lot of policy debates and environmental lawsuits regarding the effectiveness and flexibility of such regulations, and also their impacts on the premature closure of fossil-fuel generating units and the power system reliability. This chapter will provide new evidence towards these debates by studying a self-selected mixture of policy instrument using data from the Texas power generating units.

The analysis of this chapter centers on the voluntary cap-and-trade feature of the Texas regulatory framework. This feature brings strategic interactions among the regulated units and could result in different welfare outcomes compared to a single mandatory cap-and-trade framework. Which kind of policy framework is more cost-effective will rely on the distributions of abatement costs and targets across all the units. Under the Texas regulatory framework, the cap-and-trade program and intensity standards impose distinct cost structures on the regulated units. Because the cap-and-trade market equilibrium depends on the set of the participants, the post-1971 units will only volunteer for it if their abatement costs are much lower than other participants. With poor enforcement of the intensity standards, some low-cost post-1971 units might find out that they could pay lower abatement costs if they stay outside of the cap-and-trade and violate the intensity standards. On the contrary, if these units participate the cap-and-trade market, they are able to cut more emissions and sell the extra allowances at lower prices to reduce the total abatement costs. It implies that replacing the mixed policy framework by a mandatory cap-and-trade regulation could lead to higher aggregate welfare, despite the profit loss by a sub-group of units. Furthermore, although the cap-and-trade program is mandatory for the pre-1971 units, they could choose to exit from the production market if the compliance costs are too high for them. The exit decisions will further change the cap-and-trade market equilibrium and strategically interact with the volunteer decisions of post-1971 units. Without the information of each unit's abatement cost, it is impossible to predict the efficacy of the mixed regulatory framework in Texas. Since the individual abatement costs are publicly unavailable, this chapter will answer the research questions with a structural approach to recover the individual abatement costs and conduct counterfactual experiments.

The empirical strategy of this chapter starts with a structural model to link the compliance choices of generating units with their abatement costs under the existing regulations. Because the pre-1971 units have to make a one-time volunteer decision, I develop a two-stage static model to specify units' payoffs and characterize how they make the policy instrument choices. In the first stage of the model, units play a strategic game where they make the binary cap-and-trade participation decisions. Pre-1971 units could choose to exit the market in this stage and receive the scrap value. Then, in the second stage, operating units select the type of abatement technology and choose their quantity of production to maximize their payoffs. The payoffs include the revenue of production and the operating and abatement costs. The abatement costs depend on not only their own technology and production choices in the second stage but also the policy instrument choices in the first stage. Units' policy instrument choices will also strategically interact with each other by affecting the equilibrium emission price in the cap-and-trade market. The model also includes an unobserved and persistent cost shock of each unit regardless of their policy instrument choice, in order to capture the abatement cost heterogeneity across all the units.

The identification strategies to recover individual cost parameters make use of the two-stage nature of the model. Intuitively, given the set of cap-and-trade participants and the cost structure imposed by the policy instruments, the abatement cost heterogeneity among the generating units induce them to choose different levels of production and emission abatement technology in the second stage. I exploit the variations of unit-level production, abatement technology investment to identify operating and abatement costs. To deal with the endogeneity problem that these compliance choices are correlated with the unobserved cost heterogeneity, I construct instrument variables using the choices of other units in the same firm. The relevance assumption is that without the unobserved heterogeneity, all units in the same firm tend to make the same choices. The exclusion restriction is based on the fact that the unobserved abatement cost heterogeneity mainly affects the fuel combustion process where the formation of NO_x occurs. Because each unit has an independent fuel boiler, I assume the unobserved heterogeneity is uncorrelated within a firm. The estimated individual abatement technology costs are similar to the industry average costs estimated by the EPA, which indicates the instrument variables are well-constructed. Still, there are unknown parameters like the scrap values of the exit units, which are not directly correlated with the choices in the second stage. The next step is to go back to the first stage of the structural model to construct the participation game. The equilibrium strategies of exit and volunteer decisions enable the identification of the rest of the unknown parameters.

Given all the estimated parameters, I conduct counterfactual simulations to compare the outcomes under the mixed policy framework with the mandatory cap-and-trade regulation. Counterfactual analyses show that the mandatory cap-and-trade regulation aggregately outperforms the self-selected mixture of policy instruments. The overall welfare gain by implementing the mandatory cap-and-trade regulation in place of the mixed policy instruments ranges from \$176m to \$549m per year (8% to 25% relative to the aggregate welfare of the mixed framework). The exact value depends on the number of emission markets and whether the cap-and-trade policy and intensity standards are overlapping with each other. The welfare measure accounts for both the reduction of NO_x emission damage to human health and the increase in industry profits. Regarding the effects of the policy instruments on market exit, I find that the mandatory cap-and-trade regulation reduces the exit rate of pre-1971 units by around 14% compared to the mixed policy framework.

In the counterfactual simulations, I also investigate the distributional effects of the changes in policy instruments. I find that low-cost and large units owned by a few utility firms receive most of the profit increase under mandatory cap-and-trade regulation. It indicates that the existing mixed policy framework would bring benefits to the majority of generating units with small capacity and high abatement costs. Moreover, I document that under the mandatory cap-and-trade regulation, adopting a single emission market increases the emissions generated from counties in the east Texas compared to the regime of dividing into separated emission markets. Such increases in emissions could present a threat to the air quality attainment in these counties and might cause serious health problems on the local population. It implies a mandatory cap-and-trade regulation with separated emission markets could be more effective at achieving the air quality attainment in the east Texas.

This chapter is related to several areas of the literature about differentiated environmental regulations. The main contribution is presenting a new framework for welfare evaluation on unit-level self-selected mixture of environmental policy instruments, which is under-addressed in the literature. In the chapter, the decisions of market exit, policy instruments, abatement technology investment and production are integrated all together into a single structural model. Motivated by the static feature of the regulations being studied, I incorporate the exit choice into a static model, which makes this chapter methodologically different from the literature on evaluating the effects of regulation designs on the entry and exit using dynamic models (Dardati 2016; Fowlie, Reguant and Ryan 2016). Moreover, the identification strategies make full use of public available data to recover individual abatement costs, and the estimation results fit well with industry average engineering estimates. Given the estimates of cost parameters, the model is flexible enough to predict equilibrium outcomes under a wide range of regulatory regimes, and the counterfactual experiments can offer important policy implications to states and countries where there are debates over differentiated or mixed regulations.

In the literature on the choices of environmental policy instruments, Goulder and Parry (2008) and Lehmann (2012) give comprehensive reviews of the existing literature on mandatory policies as well as the use of multiple policies. This chapter is most closely related to the discussions about the choices between uniform cap-and-trade and other alternative mixed or differentiated policies. The most relevant paper is Bushnell et. al (2015), which applies the theoretical framework of Fischer (2003) to simulate the effects of a hypothetical cap-and-trade coalition among western states under the proposed Clean Power Plan. They argue that the mixed regulations on different states lead to inefficacy because the states fail to coordinate through the production market. Another related paper is Fowlie and Muller (2013) that compares uniform regulations with differentiated regulations based on non-uniform spatial emission damage. Although these papers compare a variety of uniform or differentiated policy instruments, in their models the choices of policy instruments are pre-determined by the states or the regulator. On the contrary, this chapter explicitly model and assess the welfare outcomes of the self-selection of policy instruments at the generating unit level. The model also allows me to examine the effects of policy instruments on unit closure, which is neglected by the previous papers. Moreover, most papers in the literature use industry average or engineering costs for welfare evaluation, but this chapter recovers all the relevant economic costs at the unit level by estimating the structural model.

This chapter also relates to the literature on voluntary environmental programs. Previous studies (Montero (2000); Kerr and Van Benthem (2010);

Bushnell (2011); Millard-Ball (2013)) center on theoretically addressing the optimal allowance allocation rule to induce low-cost units to volunteer for the programs. Montero (1999) is an empirical study that compares the emissions of the volunteers before and after participation in the first phase of the Acid Rain Program. He finds the evidence of adverse selection that the volunteer units are more likely to be those with excess allowances but not necessarily lower abatement costs. Contrary to the substantial emphasize on the asymmetric information between the regulator and volunteers in the literature, I assess the ex-post benefits of the voluntary provision by developing a novel structural model to characterize units' choices in various dimensions. The model does not assume the optimality of policy design or the distribution function of unit types. My welfare evaluation emphasizes the comparison of the welfare outcomes between voluntary and mandatory regulations.

Additionally, the literature has detailed examinations of the grandfathering provisions that exempted existing units in traditional vintage-differentiated regulations. The literature documents that these regulations result in both costly entry barriers and emission leakage from the emission-intensive incumbents (Nelson et al. 1993; Santon 1993; Heutel 2011; Bushnell and Wolfram 2012). However, the regulations studied in this chapter are the opposite of these traditional grandfathering provisions, as the older generation of units face with a mandatory cap-and-trade regulation while the newer generations could choose between cap-and-trade and intensity standards. This chapter adds to the literature by evaluating the effects of vintage-differentiated choices of policy instruments and comparing with a uniform cap-and-trade instrument.

The welfare evaluation of this chapter also has limitations due to the limited scopes of the regulations being studied. The regulations target on a subset of generating units in the wholesale market and exempted new entrants after the year 2000. After the year 2005 more stringent federal regulations came into effect so that those regulations discussed in this chapter were effective for a relatively short period of time. Therefore, I will not address the effects of the regulations on electricity wholesale market equilibrium and consumer surplus, and I will not examine the impacts of the mixed policy framework on the industry entry and exit in the long run.

2.2 Data and Descriptive Evidence

The data used in this chapter were obtained from different sources. The Texas Commission on Environmental Commission provided me with the allowance allocation, transaction and emission data of the cap-and-trade program. The emission data of units not participating in the cap-and-trade program were taken from the EPA air market program database. Data on unit-month level production, abatement technology choice, and other characteristics were obtained from the Energy Information Administration's forms (EIA-767/860, 923, 423/920/926).

The electricity price data used in this paper were the annual average peak-hour prices of each wholesale market from the yearly reports of the Federal Energy Regulatory Commission. All the wholesale markets with Texas generating units have active short-term bilateral transactions of electricity, which accounts for the majority of total electricity supply in these markets. Usually the on-peak hours cover 7am to 10pm on Monday to Friday. The electricity price used here is the annual volume-weighted price of the bilateral transactions for electricity delivered during the on-peak hours. This is no other better available source of data for electricity price in all the wholesale markets. I will use this electricity price measure to quantify the revenue of electricity production.

Table 2.1: Average Day Ahead Bilateral Transaction Price ($/MWh)

	2001	2002	2003	2004
ERCOT	30.21	34.98	46.90	55.2
South Central	33.88	32.33	42.84	52.61
Southwest	-	38.38	49.43	55.86
Southeast	34.21	32.84	42.90	53.82

Note: This table shows the average day ahead bilateral transaction price of electricity in each control period. The original annual prices were taken from the yearly reports of the Federal Energy Regulatory Commission. The prices shown in this table are calculated as the weighted average of the two years' price in the same control period. A control period is defined as from May of one year to the April of the next year. For Southwest NERC region, because the daily prices were intermittent in the year 2001, there was no reliable source of annual average price data.

The sample period is from the May of 2001 to the April of 2005, which consists of four control periods starting from the time the volunteer decisions were made to two years after the start of the cap-and-trade program. Units might make abatement technology investment before the start of the cap-and-trade market. This sample period enables me to observe units' production and abatement technology choices before and after their changes in compliance decisions, which is useful for the estimation of variable costs and the performance of abatement technologies in the following sections. The original monthly data from EIA were aggregated into the corresponding control period. The electricity price of each control period is also calculated as the weighted average of the two years' price in the same control period. Table 2.1 shows the price data used in this chapter. The sample period did not include the following years because the EIA did not collect unit level data in 2006 and also after 2006 there were new federal regulations introduced. Within this period, units' compliance decisions were only affected by the regulations discussed in the previous section. As the sample period is relatively short, it is reasonable to assume that the cost of abatement does not change with time.

The sample of units used in this chapter include all the generating units that had already started operation in 2000. Units retired before 2000 or entered into the electricity production market after 2000 were not included. Because the owners of the new entrants are often different from the units in the sample, and because the new entrants are exempted from regulations mentioned previously, this chapter will not address how the regulations affect production market entry. Appendix gives the detailed description of the data cleaning process.

There are many abatement technologies for generating units to reduce NO_x emissions. Consistent with the conventions in the literature, I aggregate the technologies into three categories based on the mechanism of emission reduction and the performance. The first one is referred as the Combustion

Modification in the rest of the chapter, which includes the methods of injecting air and water or adjusting combustion temperature to reduce NO_x formation during the fuel combustion process of electricity generation. Another category is referred as the Lower NO_x Burner, which is a special type of boiler equipment to control the fuel combustion process to reduce NO_x formation. The last category is the Selective Catalyst Reduction that uses catalyst to absorb NO_x after the combustion process.[1] The Selective Catalyst Reduction could achieve 80%-90% reduction on average but with very high capital cost. The other two usually achieve 40%-50% reduction and have relatively lower capital cost. Using these technologies also increases the marginal costs of production by various degrees. Appendix provides the range of industry average fixed and variable costs estimates by the EPA in 1999. Unfortunately, the cost data on each individual unit is unavailable. This chapter will recover these costs by estimating the structural model.

Table 2.2 presents the summary statistics of the characteristics of sample generating units. There are 205 units in the sample and two-thirds of them are located in east Texas. I divided the units into four groups based on their location and vintage. The majority the sample units are pre-1971 units, and

[1]The Selective Catalyst Reduction can be used together with the other two categories, but the performance of the combination is approximately the same as using the Selective Catalyst Reduction alone (see Appendix). Therefore, I treat the choice of combing the Selective Catalyst Reduction with any other combustion process control technologies as the choice of the Selective Catalyst Reduction alone. There are also other categories of technology available in the industry practice in addition to the three categories mentioned in this chapter, but they are not included in the choice set here because none of the units installed them during the sample period.

Table 2.2: Summary Statistics of Units Characteristics by Region

Variable	Mean	Std. Dev.	Min	Max
East Pre-1971 Units (Obs=105)				
Exit units	0.210	0.409	0	1
Natural Gas Dummy	0.962	0.192	0	1
Capacity (MW)	268	196	31.2	799.2
Initital Service Year	1963	8	1945	1975
Cap-and-trade Allowances (tons)	556	689	0	3103
Houston-Galveston Dummy	0.181	0.387	0	1
Dallas-Fort Worth Dummy	0.248	0.434	0	1
East and Central Texas Dummy	0.533	0.501	0	1
Beaumont-Port Arthur Dummy	0.038	0.192	0	1
East Post-1971 Units (Obs=45)				
Volunteer units	0.178	0.387	0	1
Natural Gas Dummy	0.422	0.499	0	1
Capacity (MW)	450	265	25	813.4
Initital Service Year	1981	9	1954	1998
Cap-and-trade Allowances (tons)	2008	1765	15	5643
Houston-Galveston Dummy	0.244	0.435	0	1
Dallas-Fort Worth Dummy	0.089	0.288	0	1
East and Central Texas Dummy	0.644	0.484	0	1
Beaumont-Port Arthur Dummy	0.022	0.149	0	1
West Pre-1971 Units (Obs=42)				
Exit units	0.238	0.431	0	1
Natural Gas Dummy	1.000	0.000	1	1
Capacity (MW)	135	129	18.4	535.5
Initital Service Year	1961	8	1947	1974
Cap-and-trade Allowances (tons)	446	539	0	2575
West Post-1971 Units (Obs=13)				
Volunteer units	0.077	0.277	0	1
Natural Gas Dummy	0.538	0.519	0	1
Capacity (MW)	298	213	53.6	720
Initial Service Year	1981	7	1975	1994
Cap-and-trade Allowances (tons)	3150	4243	25	14794

Note: In each panel, the row 'Exit units' shows the summary statistics of the dummy variable indicating whether the pre-1971 units exited the market in or before 2004. The row 'Volunteer units' shows the summary statistics of the dummy variable indicating whether the post-1971 units is a cap-and-trade volunteer unit. The 'Initial Service Year' row is the average initial year of commercial service of the units in each region. The 'Cap-and-trade Allowances' row shows the actual allowances allocated to cap-and-trade participants (including the allocation to exit units) and the potential allowances the outside post-1971 units would have if they are included in the cap-and-trade program.

Table 2.3: Summary Statistics of Unit Production and Emission

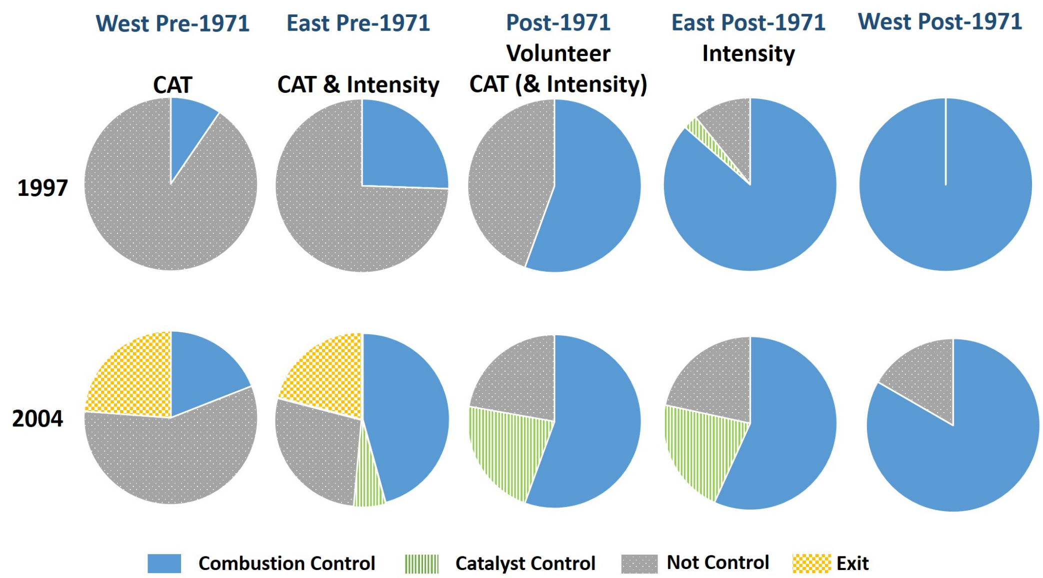
Variable	Obs	Mean	Std. Dev.	Min	Max
heat input (1000,000 MMBtu)	737	11.929	16.780	0.000	67.679
generation (1000,000 MWh)	737	1.155	1.660	0	7
efficiency (MWh/MMBtu)	737	0.089	0.017	0	0.239
emission (tons)	737	1162	1953	0	9316
Combustion Modification Dummy	737	0.300	0.459	0	1
Lower NOx Burner Dummy	737	0.213	0.410	0	1
Selective Catalyst Reduction Dummy	737	0.061	0.240	0	1

Note: This table shows the summary statistics of unit-level production and emission choices during May 2001- April 2005. Each observation is at the unit/control period level. A control period is defined as from May of one year to the April of the next year. Heat input is measured in the unit of 1000,000 MMBtu. Generation is measured in the unit of 1000,000 MWh. Efficiency is calculated as the ratio between generation and heat input per unit per control period. The last three rows show the fraction of observations choosing each category of abatement technology. Observations for exit units are dropped.

around 22% of them exit the market during the sample period. The volunteers are really rare, as there are only 9 volunteering units altogether. Newer and larger units are concentrated in the east Texas. Over 60% of east generating units are located in the East and Central Texas region. In the west, all pre-1971 units are natural gas units, and they are of the lowest capacity and oldest vintage. Most coal units belong to the post-1971 units located in the east Texas.

Table 2.3 presents the summary statistics of annual unit level heat input, generation and emission. To show that the abatement choices are correlated with the regulatory policy instrument, Figure 2.1 compares the abatement technology choices of different unit groups before and after the regulations were introduced. The units are divided into five groups: west pre-1971 units, east pre-1971 units, post-1971 units volunteering for the cap-and-trade program, and post-1971 units outside of cap-and-trade in both east and west Texas. In each pie of Figure 2.1, the fraction with the gray color represents the fraction of units without any abatement technology installed. The blue fraction represents the units using the Combustion Modification or Lower NO_x Burner. The green fraction includes the units using the Selective Catalyst Reduction. In the first pie of the second row, the yellow fraction represents the retired pre-1971 units in 2004. This graph shows that the pre-1971 units complied with the cap-and-trade regulations by retirement or using the inferior technologies. A higher fraction of pre-1971 units in east install abatement technologies compared to those in the west. It might be caused by the extra intensity standards in the east Texas or the possibility that east units have lower abatement costs compared to the west units. A large proportion of post-1971 units chose to install the best technology to comply with the stringent intensity standards. The post-1971 units volunteer for the cap-and-trade program could also sell extra allowances to the pre-1971 units by reducing their own emissions with the best abatement technology. In the west Texas, no post-1971 units installed the best technology, which might be caused by the fact that there was no extra abatement requirement relative to the year 1997 on these units. This graph shows that the abatement technology choice was correlated with the vintage and the regulation instrument. The fact that the pre-1971 units were unlikely to install the catalyst reduction technology might indicate that the abatement costs of the pre-1971 units were relatively higher. As discussed in Section 2, this suggests that limiting the cap-and-trade program to the pre-1971 units might not be cost-minimizing.

Figure 2.1: Units Abatement Choice before and after the Regulations



Note: The first row shows the units' choices in the base year 1997 and the second row shows the choices in the control period 2004. The gray color represents the fraction of units without any abatement technology; the blue color represents the fraction of units with combustion process abatement technology (Combustion Modification or Lower NO_x Burner); the green color represents the fraction of units with the Selective Catalyst Reduction. The orange color represents retired pre-1971 units.

The previous sections also mention that a difference between cap-and-trade and intensity regulation is that the cap-and-trade regulation binds the absolute amount of emissions, indicating that the cap-and-trade participants are more likely to comply by cutting production. To illustrate this difference, Table 2.4 reports the results from a first-difference regression that regresses the changes of heat input between 2004 and 1997 on each individual unit cap-and-trade participation status in 2004. A dummy variable for the units subject to the intensity regulation in the east Texas is also included in the regression. Thus, the baseline group consists of the units in the west Texas but not in the cap-and-trade program. Table 2.4 shows that controlling for unit fixed effect, the cap-and-trade participants are more likely to reduce their heat input relative to 1997, but there is no significant reduction for the units subject to the intensity standards. It suggests that the abatement choices of whether cutting production are correlated with the type of environmental policy instrument.

Table 2.4: Production Reduction in Cap and Trade

	heat input 2004-heat input 1997
cap-and-trade	-2.963***
	(1.039)
east intensity regulation	1.455
	(0.933)
sample selection	12.143
	(8.421)
constant	-4.970*
	(2.800)
Observations	187
R-squared	0.058

Note: This table presents the estimation results of the effects of the changes in regulation on heat input change between 2004 and 1997. The results are measured in 1,000,000MMBtu. The sample selection variable is the inverse Mills ratio to control to sample selection problem caused by units retirement and the regression results are estimated via the Heckit two-step method. The first step is to run the probit regression on the retirement decision on unit characteristics such as generator capacity, initial service year, wholesale market region and so on. Then the retirement probability is estimated for each unit. The second step is to add a variable which is the inverse Mills ratio of the predict retirement probability into the regression, denoted by 'sample selection' in this table. Bootstrap standard errors are in the parentheses (*** p¡0.01, ** p¡0.05, * p¡0.1).

The descriptive evidence in this section shows that units make different abatement technology and production choices under different regulation instruments, and that the pre-1971 and post-1971 units seem to have different abatement costs. Given different policy instruments impose different cost structures on the generating units, it is necessary to develop a model to link the abatement costs of units with their abatement and production choices and also the regulator framework. The next section will present the model.

2.3 Model

This section presents a structural model of generating unit abatement compliance behavior taking the existing policies as given. The model is to recover the abatement costs of generating units and simulate for equilibrium outcomes under a counterfactual cap-and-trade regulation. The model does not impose any assumption on the optimality of existing policies or the regulator's objective function. In the model, the decision makers are the individual profit-maximizing units since the as the regulator allows each unit to have its own specific policy instrument and compliance choice.

The model is a two-stage, static model. Because the volunteer decision has to be made before the start of the cap-and-trade program and cannot be changed later on, there are very limited dynamic interactions between the generating units. Therefore, I use a static model to characterize how units make compliance decisions under the existing policy framework. I also assume that units' payoffs do not vary significantly from year to year. They only need to make a one-time choice of the policy instrument. Once the units make the decisions, they cannot change it. The timing of the model is the following:

1. In the first stage, the regulations are announced to all units and each unit i decides whether to participate in the cap-and-trade program by choosing $d_i \in \{0, 1\}$, where $d_i = 1$ represents participation. If a pre-1971 unit chooses not to participate, it will be regarded as exiting from the production market. The exit unit will only receive its scrap value, which will be specified later.

2. In the second stage, the participation decisions in the first stage are revealed to all the units. Cap and trade participating units ($d_i = 1$) form the emission markets. Based on the regulation background, there are two markets: East and West Texas. Despite the participation decisions, all the units choose the type of abatement technology to install ($j_i \in \{none, CM, LNB, SCR\}$), and then choose the amount of production (heat input h_i) for a control year. Then, the units sell the electricity and pay for the associated operating and abatement costs. The cap-and-trade policy and the intensity standards also have differential impacts on units' payoffs in this stage. The payoff functions and the choice rules of technology and heat input will be specified later.

There are several maintained assumptions in the model. First of all, I assume the price of electricity is constant. The sample of units in this chapter is a subset of all electric generating units in the production market, and the total capacity of the market is always higher than the peak demand. Therefore, the sample of units could not have substantial effects on the production market price. In this model, units make production and emission decisions for a whole year, and I assume that units have perfect information to predict annual average price without considering the day by day variations of prices in the wholesale market. The units will take the annual average electricity price as given to choose the production level. Meanwhile, there is concern that the fossil-fuel units are the price-setting units based on the merit order of the wholesale electricity market. Appendix will show the results are robust when relaxing this assumption.

Another assumption is that the units do not choose their emissions directly. They choose the type of abatement technology and their heat input. Emission is a function of individual fixed effect, the type of abatement technology and the amount of heat input. The specification of the emission e^i_j by unit i with abatement technology j is the following:

$$e^i_j = a_i exp(\sum_j \gamma_j \cdot 1(techj_i = 1))h_i, \tag{2.1}$$

where a_i is unit is emission fixed effect, γ_j is the parameter to represent the performance of the technology j, and $techj_i$ is the dummy variable indicating whether unit i chooses the optimal technology j_i. Given the specification of the emission function, the emission intensity is defined as

$$r^i_j = e^i_j / h_i.$$

The first stage of the model is assumed to be a complete information game where units know everyone's costs[2]. Since the regulations are also public available information, it implies that the payoffs of every unit for every possible $d = (d_1, ..., d_N)$ is also common knowledge. The units are forward-looking in the first stage. They hold correct beliefs about the emission price under the cap-and-trade regulation and they know everyone's optimal abatement investment and production conditional on each d in the second stage. Therefore, the specifications of payoffs below will begin with the second stage of the model and then go back to the first stage.

2.3.1 Second Stage Payoffs

In the second stage, the production decision is modeled as the choice of heat input. The production function is specified as $Q_i = q_i h_i$, where q_i is the unit i's fuel efficiency factor and Q_i is its production. This chapter assumes each unit has a constant fuel efficiency that is unaffected by the choice of abatement technology, which is also a public information. This assumption is empirically supported by the regression analysis in Appendix. Units receive production revenue pQ_i with electricity price p. Each wholesale market has its own specific electricity price. To make the notation tractable, here I do not explicitly specify the index to represent the wholesale market.

[2]Hortacsu and Puller (2008) which studies the ERCOT spot market argues that firms in ERCOT typically know others' marginal costs of generation. Because units in ERCOT accounts for the majority of Texas generating units, this chapter assumes that there is no private information about costs

The next part of the payoff is unit i's variable cost, which depends on its own choice of the abatement technology and all units' choices of the regulation instruments. It is specified as below:

$$C(d, j_i, h_i) = (c_1 h_i + c_2 h_i^2) + [\sum \beta_j \cdot 1(techj_i = 1) + u_i] \cdot h_i + d_i \cdot \tau(d, A) \cdot (e_j^i - A_i)$$
$$+ \lambda \cdot Violate \cdot Q_i \cdot (r_j^i / Target_i - 1). \tag{2.2}$$

The expression $(c_1 h_i + c_2 h_i^2)$ is the operational cost of fuel and labor, where both c_1 and c_2 are assumed to be positive. This nonlinear assumption is to capture the fact that units usually do not operate at their full capacity. With increasing level of production, the units may face with increasing risks of broken down and have to pay extra money for maintenance and transmission congestion, which suggests that the operating cost of production is convex. The second part of the costs $[\sum \beta_j \cdot 1(tech_j^i = 1) + u_i] \cdot h_i$ is the abatement technology variable cost. This part of the cost is assumed to be linear in heat input and the cost parameter β_j varies with the category of technology. The u_i is each units abatement cost shock which is unobservable to the econometrician, but is the public information for all the units. This shock depends on the characteristics of the fuel combustion process unobserved from the data, such as the temperature and residence time of fuel burning. It varies from unit to unit since each unit has a separate fuel boiler. Also, it does not vary with the choice of regulation instrument.

The rest of the variable cost is determined by the regulation instruments and the cap-and-trade participation decisions of all units. In the cap-and-trade program, the unit i faces the emission price τ and the associated emission cost of buying allowances from others $\tau(d, A) \cdot (e_j^i - A_i)$, where A_i is the individual allowance allocation. The emission price $\tau(d, A)$ is assumed to be the market clear price given all units' participation decisions and the total allowance in the market. In the second stage, the participation decisions and the allocation rule are revealed so that there is a constant price in each market. The units in each cap-and-trade market take the price as given to choose the abatement technology and heat input and no single unit is able to alter the price. For the units subject to the intensity regulation, they face the variable penalty tax $\lambda \cdot Violate \cdot Q_j^i \cdot (r_j^i / Target_i - 1)$ in case of violation. The tax rate is λ and the tax is proportional to the production and the degree of violation as discussed in the previous chapter.

For each unit i, given its choice of abatement technology and heat input, its variable profits without accounting for the fixed retrofit cost of abatement technology is set to be

$$\bar{\pi}(d, j_i, h_i) = pq_i h_i - C(d, j_i, h_i). \tag{2.3}$$

Assuming that units choose the abatement technology j_i before choosing their heat input h_i, the optimal level of heat input maximizing the variable profits varies with the type of technology. Therefore, given the participation decisions d, for the technology j_i the optimal heat input $h^*(d, j_i)$ solves the FOC as the following:

$$\frac{\partial \bar{\pi}(d, j_i, h_i)}{\partial h_i} = 0. \tag{2.4}$$

After solving for the heat input choice, we can solve for each unit's variable profits given its choice of the abatement technology, which is denoted as $\bar{\pi}(d, j_i, h^*(d, j_i))$. As mentioned above, there are additional fixed costs of installation and retrofit for every category of abatement technology. It is specified as the following:

$$F(j_i) = F_{j0} + F_{j1} \cdot InitialServiceYear_i + F_{j2} \cdot Capacity_i - \varepsilon_{ij}. \tag{2.5}$$

The ε_{ij} is the unobserved fixed cost shock to the econometrician, and it is assumed to follow Type I extreme value distribution. It is also public information among the units. Its scale is normalized to be \$3m since the fixed cost is usually very high. Because the industry groups always argue about how costly it is for old and small units to retrofit in order to suspend the introduction of new regulations, the specification includes interaction terms with the initial service year and the capacity to quantify their effects on the retrofit fixed costs. For the choice of abatement technology, I assume that the units choose the one that maximizes the difference between their variable profits and the fixed costs:

$$\pi(d, j_i, h_i) = \bar{\pi}(d, j_i, h^*(d, j_i)) - F(j_i) - 1/2 \cdot Violate \cdot (F(SCR_i) - F(j_i)) + \varepsilon_{ij}. \tag{2.6}$$

The profits in equation 2.6 also includes the intensity regulation penalty on fixed costs. Based on the aggregation of abatement technologies in the previous section, the violating units will be penalized for half of the cost difference between their chosen technology and the SCR technology. The choice of the abatement technology is a discrete choice problem, and unit i will choose the profit maximizing technology $j^*(d)$ given the cap-and-trade participation decision. Therefore, the payoff of unit i in the second stage of the model is denoted as $\pi(d_i, j^*(d), h^*(d, j^*(d)))$. If in the first stage of the model a pre-1971 unit decides to exit, then in the second stage it receives 0. This also implies that even if an exit unit receives non-zero allowances, these allowances have no value.

2.3.2 First Stage Payoffs

The first stage is modeled as a normal form game where each individual faces a binary choice of whether to take part in the cap-and-trade program. The model also assumes that every unit i receives a payment S_i in additional to its payoff π from the second stage when not participating in the cap-and-trade program. For the pre-1971 units the S_i represents the scrap value of retirement, and for the post-1971 units the S_i can be interpreted as the benefits of the intensity regulation under which there is more flexibility to expand production. Since the abatement targets imposed by the regulators often depend on the historical emissions and heat input, when future regulation uses the current period as the base year to define abatement targets, choosing to not volunteer for the cap-and-trade suggests the potential production expansion might give the units some additional benefits. For post-1971 units, the S represents their beliefs on such future benefits. In the rest of the chapter, the S_i will be referred as the outside benefits. S_i for unit i is specified as the following:

$$S_i = \theta_{0i} + \theta_{1i} \cdot east_i + \theta_{2i} \cdot post1971_i + \theta_{3i} \cdot east_i \cdot post1971_i,$$

$$\theta_{ki} = \mu_k + \sigma_k \cdot v_i, v_i \sim N(0,1)iid, k = 0, ..., 3.$$

The random coefficient specification is to allow each individual unit to have its own specific outside benefit. Its value depends on not only the emission market but also the vintage, which is because the interpretation differs across the vintage. In the first stage, the nature independently draws the v_i from the standard normal distribution for each unit i and announces its value to all the units. Thus, the values of S of all the units are public information. Units also hold beliefs that in the second stage, every unit will make the optimal choice of abatement technology and associated heat input based on the choice rule specified in the previous subsection. They also know the costs of all the units and hold correct beliefs about the market clear emission price in each configuration of the cap-and-trade emission market in the second stage. There is no private information in this game.

The strategy for each unit i in the first stage is defined as the binary choice of participating in the cap-and-trade program ($d_i \in \{0, 1\}$). Units could also play mixed strategies. This chapter assumes each unit makes such choice independently, as the Texas regulator allows each unit to make an independent choice in the real policy. If unit i chooses cap-and-trade, its total payoffs will be

$$\pi((1, d_{-i}), j_i^*(1, d_{-i}), h_i^*[(1, d_{-i}), j_i^*(1, d_{-i})])$$

where $j_i^*(1, d_{-i})$ and $h_i^*[(1, d_{-i}), j_i^*(1, d_{-i})]$ will be the associated optimal choices of technology and heat input in the cap-and-trade program. Theses choices will also depend on the participation choices of other units, since the equilibrium allowances prices depend on who are in the emission market. If unit i chooses $d_i = 0$, its payoff will be $\pi(0, j_i^{**}(0), h_i^{**}[0, j_i^{**}(0)]) + S_i$. This payoff is no longer affected by other units, and the associated optimal technology and heat input could be different from the optimal choices in the cap-and-trade market.

I have specified all the components of units' payoffs and how they make the optimal choices in each stage of the model. Suppose a emission market has N units to make the cap-and-trade participation decision. The definition of the equilibrium in this market is as the following:

The equilibrium is a list of choice variables $(d_i, j_i, h_i, e_i)_{i=1}^N$ *and equilibrium emission prices* $\tau(\{d_i\}_{i=1}^N, \{A_i\}_{i=1}^N)$ *such that*

1. *For any unit i,* h_i *maximizes its variable profits in the second stage given the choices of* $(\{d_i\}_{i=1}^N, j_i)$ *and* $\tau(\{d_i\}_{i=1}^N, \{A_i\}_{i=1}^N)$ *in the cap-and-trade emission market. If unit i is not regulated by cap-and-trade, it just takes its own intensity targets as given to choose the optimal heat input. This implies that* h_i *is the solution of* $\frac{\partial \bar{\pi}(\{d_i\}_{i=1}^N, j_i, h_i)}{\partial h_i} = 0$.

2. *For any unit i, j_i maximizes its second stage payoffs $\pi(\{d_i\}_{i=1}^N, j_i, h_i)$ given $\{d_i\}_{i=1}^N$ and $\tau(\{d_i\}_{i=1}^N, \{A_i\}_{i=1}^N)$ in the cap-and-trade emission market. Again, If unit i is not regulated by cap-and-trade, it just takes its own intensity targets as given to choose the optimal category of abatement technology to maximizes $\pi(d_i = 0, j_i, h_i)$.*

3. *For any unit i, $d_i = \underset{d_i \in \{1,0\}}{argmax} \{\pi((d_i = 1, d_{-i}), j_i^*, h_i^*), \pi(d_i = 0, j_i^{**}, h_i^{**}) + S_i\}$, where (j_i^*, h_i^*) are the solutions specified in parts 1 and 2 given $(d_i = 1, d_{-i})$, and (j_i^{**}, h_i^{**}) are the solutions specified in parts 1 and 2 given $d_i = 0$.*

4. *In the cap-and-trade market, total emissions equal to total allowances such that $\sum_{i=1}^N d_i \cdot e_i = \sum_{i=1}^N d_i \cdot A_i$, where for any unit i, its emission is specified as $e_j^i = a_i exp(\sum_j \gamma_j \cdot 1(techj_i = 1))h_i$. Here h_i is the optimal heat input specified above and $techj_i$ is the dummy variable indicating whether unit i chooses the optimal technology j_i as specified above.*

Each emission market will have its own specific equilibrium. Because the number of units and the dimension of choices are finite, the Nash equilibrium always exists. This chapter also assumes that units' payoffs do not vary significantly from year to year. They only need to consider one-period payoff to make the participation choice. [3]

[3] Although the data sample consists of multiple years of data, this is to enable me to observe units' production and abatement technology choices before and after their changes in compliance choices, which is helpful for the estimation of variable costs and the performance of abatement technologies in the next section. The sample of data does not include

2.4 Identification and Estimation

2.4.1 Identification Strategies

In this model, the unknown parameters include the abatement technology reduction factor γ, the operating cost c, the variable and fixed abatement technology cost β and F, the penalty tax rate λ, and outside benefits S. This subsection will specify the identification strategies for these parameters step by step.

The abatement technology reduction factor γ can be identified based on the specification of emission function and the changes of abatement technology choice by generating units during the sample period. The emission function is specified as in equation 2.1. Taking the log on both sides we can obtain

$$log(e_i) - log(h_i) = \gamma_1 CM_{it} + \gamma_2 LNB_i + \gamma_3 SCR_i + loga_i. \qquad (2.7)$$

Then the γs are estimated by running a unit fixed effect regression with $loga_i$ being the fixed effect with multiple periods of data.

For the variable costs and penalty tax rate (c, β, λ), the identification relies on the optimal choice rule of heat input. Recall that the optimal heat input maximizes variable profits conditional on the choice of cap-and-trade participation and abatement technology as given in equation 2.4, which can

observations for a long period of time after the cap-and-trade program starts. Due to this data limitation, I am unable to consider units' dynamic choices in the long run. Therefore, the model is static. For units' cap-and-trade participation choices, I only consider their status in the control period 2004.

be specified as the following:

$$pq_i - \tau d_i \cdot \frac{e_i}{h_i} = c_1 + 2c_2 h_i + \lambda \cdot Violate \cdot q_i \cdot (\frac{r_j^i}{Target_i} - 1) + [\sum \beta_j \cdot 1(techj_i = 1) + u_i]. \tag{2.8}$$

This equation is linear in the unknown parameters $(c_1, c_2, \lambda, \beta)$, which suggests using the OLS procedure for estimation. However, the presence of the unobserved cost shock u_i causes several endogenous concerns. The higher the u_i is, the higher cost unit i has to pay for its production and emission abatement. Therefore, u_i might be negatively correlated with heat input, positively correlated with the likelihood of choosing the Combustion Modification or Lower NOx Burner technology, and also correlated with the probability of violating the intensity standards. It implies that running OLS of equation 2.8 will result in estimation bias. To deal with this problem, I construct instrument variables for each endogenous variable. The choice of the instrument variables is to supplement the individual unit's choices of production and technology by the choices of other units owned by the same utility firm or the same plant. Because decisions for the units owned by the same owner go through the same management process, without any individual abatement cost heterogeneity, it is reasonable to expect that the manager will make the same choices for all the units in the firm. As for the exclusion restriction, this chapter assumes that the cost shock is unit specific. This shock mainly depends on the fuel burning process. Since each unit has a separate fuel burning process in its fuel boiler, this chapter assumes that each unit's cost shock is uncorrelated with other units, which implies that other units' choices will be uncorrelated with each individual cost shock u_i. Thus, such instrument variables will satisfy the identification requirements.

The construction of each instrument variable is as the following. For the choice of abatement technology, the instrument is the fraction of other units adopting the same technology in the same firm. For the choice of heat input, this chapter uses the average capacity of other units in the same plant as the instrument variable assuming that higher capacity is correlated with higher production level.[4] For the identification of the violation tax rate, the instrument variable for the expression $Violate \cdot q_i \cdot (r_j^i/Target_i - 1)$ is the product of the average fuel efficiency of other units in the same firm and each unit's intensity target. The lower the target set by the regulator is, the more likely the unit is to violate the target and also the higher the degree of violation will be. Denote these instrument variables and the constant variable together as vector Z_i for unit i, the parameters in equation 2.8 is estimated via optimal weighted GMM with the moment condition $E(Z_i u_i) = 0$.

After obtaining the variable cost parameters and emission function parameters, based on the model we can construct the variable profits for each type of abatement technology $\bar{\pi}(d, j_i, h^*(d, j_i))$ for each individual unit using equation 2.3. Given the distributional assumption of the fixed cost shock ε_{ij}, we can formulate the choice probability of abatement technology for each unit,

[4]Because the average capacity of other units in the same firm is not significant in the first stage test for unit heat input, this chapter uses the average capacity of other units in the same power plant as the instrument.

and the fixed cost parameters can then be estimated via MLE.

Finally, the remaining unknown parameters are the outside benefits S. With the estimated parameters in the second stage, we can construct each unit's payoffs in and outside of the cap-and-trade program given others' participation decisions. The identification of outside benefits relies on the equilibrium strategies in the cap-and-trade participation game. Recall that the specification of the S_i for unit i is

$$S_i = \theta_{0i} + \theta_{1i} \cdot east_i + \theta_{2i} \cdot post1971_i + \theta_{3i} \cdot east_i \cdot post1971_i,$$

$$\theta_{ki} = \mu_k + \sigma_k \cdot v_i, v_i \sim N(0,1)iid, k = 0, ..., 3.$$

The unknown parameters here are $\{\mu_k, \sigma_k | k = 0, ..., 3\}$ and they will be estimated via Method of Simulated Moments. The data moments are the number of pre-1971 cap-and-trade participating units, the number of post-1971 cap-and-trade participating units, the average capacity and initial year of service of the pre-1971 participating units, and the average capacity and initial year of service of the post-1971 participating units in each emission market (see Table 2.12). The estimation begins with 50 idd draws of v for all units from the standard normal distribution. Given each guess of $\{\mu_k, \sigma_k | k = 0, ..., 3\}$, I solve the participation game for each draw of v and then construct the simulated moments. Finally, I estimate the parameters by minimizing the distance between the data moments and the simulated moments in each emission market. Each data moment is weighted by its relative scale.

This chapter uses the units' status in the control period 2004 as the participation decision in the first stage of the structural model. To reduce the dimension of the game, I also impose additional assumptions on the game players and their strategies. Appendix to this chapter gives the detailed description of the assumptions and the estimation procedure, as well as the discussion of multiple equilibria. The same procedure is also used in the counterfactual experiments in the next section. To compute the standard errors of the estimated outside benefits, I randomly draw 25 sets of parameters of the emission function, marginal and fixed costs from the estimated asymptotic distributions in previous subsections and repeat the estimation procedure for every set.

2.4.2 Estimation Results

Table 2.5 gives the estimation results of the emission function parameters (γ). The coefficients estimates imply that the Selective Catalyst Reduction has the best performance among the choice set of the technologies. The first two categories of technologies achieve approximately 30-40% reduction in emission, but the Selective Catalyst Reduction could reduce emissions by nearly 90%. Appendix provides robustness checks on alternative specifications of the emission function. With the results in Table 2.5, I will be able to predict the emissions and emission intensities of each unit using each type of the abatement technology. It will be useful for constructing units' second-stage payoffs.

For the estimation results of variable costs and penalty tax rate (c, β, λ),

Table 2.5: Emission Function Results

	log(e)-log(h)
Combustion Modification	-0.527***
	(0.075)
Lower NOx Burner	-0.410***
	(0.066)
Selective Catalyst Reduction	-1.562***
	(0.083)
Observations	737
Units	205
R-squared	0.541

Note: This table presents the results of unit-level fixed-effect regression of the equation (13) using data from 2001 to 2004 for all units with positive emissions. Observations for exit units are dropped. The constant is not reported in this table. The standard errors are in the parenthesis (*** p¡0.01, ** p¡0.05, * p¡0.1).

firstly Table 2.6 shows the first stage OLS result of regressing the potential endogenous variables on the instrument variables in equation 2.8. Table 2.6 indicates that every selected instrument is significantly correlated with the endogenous variable. The F statistics also indicate that the instrument variables are strong instruments. The estimation results of the variable cost parameters are in Table 2.7. The column (1) of Table 2.7 shows the result with the instruments and the column (2) shows the result without using the instruments. The results indicate that the operating cost accounts for a large proportion of the overall variable costs. The estimated c_2 is positively significant, which justifies the assumption of convex operating costs. The penalty tax rate is estimated to be \$3.5 per unit of production and per percentage of violation. For costs associated with abatement technology, the costs of using the Combustion Modification and the Lower NOx Burner are low and not significantly

Table 2.6: First Stage OLS

	(1) Heat Input	(2) CM	(3) LNB	(4) SCR	(5) Penalty Factor
heat input iv	0.713***	-0.002***	0.007***	0.001	0.000***
	(0.026)	(0.001)	(0.001)	(0.000)	(0.000)
CM iv	-2.664	1.092***	-0.133**	-0.008	-0.001*
	(1.682)	(0.054)	(0.052)	(0.031)	(0.001)
LNB iv	3.777*	0.099	0.912***	-0.012	-0.001
	(2.099)	(0.067)	(0.065)	(0.039)	(0.001)
SCR iv	3.290	0.166	-0.093	1.005***	0.002
	(3.830)	(0.123)	(0.119)	(0.071)	(0.002)
penalty iv	39.935***	0.598***	0.107	-0.042	-0.008***
	(4.250)	(0.137)	(0.132)	(0.079)	(0.002)
R-squared	0.556	0.387	0.341	0.244	0.102
F-stat	182.28	91.77	73.74	44.36	8.04
p-value of F	0.00	0.00	0.00	0.00	0.00

Note: This table presents the first stage OLS results of regressing the instrument on the endogenous variables of equation 2.8 using 2001-2004 data for all units with positive emission and heat input. The number of observations is 737. Observations for exit units are dropped. Standard errors are in the parentheses (*** p¡0.01, ** p¡0.05, * p¡0.1). Additional controls include year fixed effects. For cap-and-trade pre-1971 units in the west Texas their emission intensity targets are the regional emission intensities used in the allowance allocation rule in equation 1.2.1.1. Heat input is rescaled to measure in the unit of 1000,000 MMBtu. For units in the east Texas the intensity targets are taken from Table 3. For the rest of units, the target is their emission intensity in 1997. The CM, LNB, SCR are the abbreviations of Combustion Modification, Lower NOx Burner and Selective Catalyst Reduction. The last row shows the F-test statistics for the joint significance of all instrument variables.

Table 2.7: Variable Costs GMM Estimates ($/MMBtu)

	(1) IV Result	(2) No IV
c_1	3.200***	3.327***
	(0.097)	(0.060)
$c_2 * 10^6$	0.005***	0.006***
	(0.002)	(0.001)
Combustion Modification	0.226	0.319***
	(0.220)	(0.087)
Lower NOx Burner	0.139	0.207***
	(0.253)	(0.100)
Selective Catalyst Reduction	2.490***	0.827***
	(0.513)	(0.152)
intensity tax	3.586***	-2.408***
($/MWh*violation%)	(0.011)	(0.476)

Note: This table presents the estimation results of the equation (14) using 2001-2004 data for all units with positive emission and heat input. The number of observations is 737. Observations for exit units are dropped. Appendix provides additional checks on sample selection. The column (1) in this table is the result using instrument variables and the column (2) is the result without instrument variables. Standard errors are in the parentheses (*** p¡0.01, ** p¡0.05, * p¡0.1). The reported scale of the coefficient c_2 is multiplied by 10^6. All the results are measured in $/MMBtu except for the intensity tax rate.

Table 2.8: Unobserved Abatement Cost Shock and Fuel Efficiency

	Cost Shock ($/MMBtu)		Fuel Efficiency (MWh/MMBtu)	
	West	East	West	East
Pre-1971	0.034	-0.059	0.083	0.084
Post-1971 Volunteer	0.002	-0.178	0.084	0.092
Post-1971 Outside	0.117	-0.170	0.090	0.089

different from zero, but the cost of using the Selective Catalyst Reduction is pretty high and close to the scale of the operating cost. The rank order of the variable costs of abatement technologies in Table 2.7 is consistent with the industry average variable costs given by EPA (1999), which are listed in Appendix. This demonstrates that the model specification and the estimation strategies are well developed.

If we compare the results in column (1) and (2) of Table 2.7, we will find that the differences are consistent with the projected directions of estimation biases by OLS as previously mentioned in this section. Recall that the cost shock is positively correlated with the probability of choosing the Combustion Modification and the Lower NOx Burner, so that the coefficients on these two variables will be positively biased in column (2), while the coefficient on the Selective Catalyst Reduction will be negatively biased in column (2). Without the instruments, the penalty tax is negative which is possibly caused by the correlation between fuel efficiency and unobserved cost shock, but with instrument variables the estimated coefficient is positively significant. Therefore, the estimation results support the identification strategies.

With the variable costs, we can estimate the unobserved cost shock u for each unit, and investigate whether there exists adverse selection in the voluntary cap-and-trade program. Table 2.8 gives the average estimated cost shock and fuel efficiency by region and participation group. It indicates that the post-1971 volunteer units on average have the lowest cost shock compared to the other groups in each regional market, and their fuel efficiency is also

Table 2.9: Fixed Costs Estimates

	CM	LNB	SCR
Constant	84.420	361.103***	809.966***
	(106.711)	(81.110)	(169.392)
Initial Service Year	-0.040	-0.180***	-0.403***
	(0.055)	(0.041)	(0.085)
Capacity	-0.023***	-0.022***	-0.043***
	(0.003)	(0.003)	(0.005)

Note: The data used for the fixed cost estimation are the last control period data of each unit in the sample. The total observation is 205. LR Chi(9)=169.625. Results are interpreted as the cost in $m per year. In the industry practice, the Selective Catalyst Reduction usually has the life span of seven years while the other two categories can be used for longer periods of time. Standard errors the parentheses are computed via 1,500 random subsamples with subsample size 150(*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

pretty high. The results show that there is no evidence of adverse selection in this cap-and-trade program.

Table 2.10: EPA Cost Estimates 1999

	fixed ($/kW)	variable ($/MWh)
Combustion Modification	9.6-32.4	0.05-0.25
Lower NOx Burner	16.8-46.7	0-0.07
Selective Catalyst Reduction	69.7-71.8	0.24-1.27

Source: U.S. Environmental Protection Agency. Nitrogen Oxides (NOx), Why and How They Are Controlled, 1999. https://www3.epa.gov/ttncatc1/dir1/fnoxdoc.pdf. This bulletin is to provide basic guidance about the costs and mechanisms of NO_x emissions control. The rank order of the estimation results in the paper with Table 2.10 are consistent with each other.

Table 2.9 shows the estimation results of fixed costs, and it indicates that the fixed cost of the Combustion Modification is the lowest while the fixed cost of the Selective Catalyst Reduction is the highest. This rank order is also the same as the industry average given in Table 2.10. Another result of Table 2.9 is that the coefficient on the initial service year is not significant at 10% level for the Combustion Modification, suggesting the fixed costs of this technology do not differentiate across vintage. The supporters of the traditional grandfathering provision always claim that older units have to incur higher capital costs to install these technologies. The result in this chapter partially contradicts with such claim and implies that the grandfathering provision might result in regulation inefficiency. Table 2.9 also shows that units with smaller capacity have to pay higher fixed costs as all the capacity coefficients are negatively significant at 1% level, which is because the technical difficulty of adding abatement devices on small units is usually higher.

Table 2.11 presents the estimation results of the outside benefits. To make the results easier to interpret, Table 2.11 also transforms and lists the results of normal distribution parameters for outside benefits by vintage and regional market. Results show that the outside benefits are heterogeneous across different groups of units. East units on average have higher outside benefits than the west units. The benefits of not participating in the cap-and-trade program for the post-1971 units are on average higher than the scrap value of the pre-1971 units. The averages of outside benefits for all units are positive. Such extra gains make the cap-and-trade regulation more costly for the generating units, which reveals why we observe so many retired units and so few volunteers in the cap-and-trade program. Table 2.12 indicates that except for the average capacity of volunteer units, the estimation results fit well with the data.

Table 2.11: Estimation Results of S ($m)

Parameter Estimation Results							
μ_0	μ_1	μ_2	μ_3	σ_0	σ_1	σ_2	σ_3
0.009	2.794***	5.329***	-4.203***	0.126***	18.845***	4.224***	66.233***
(0.005)	(0.243)	(1.140)	(1.888)	(0.017)	(1.537)	(1.473)	(14.264)

The Normal Distribution Parameters of Estimated Outside Benefits by Region-Vintage		
unit group	mean	std
West Pre-1971	0.009	0.126
West Post-1971	5.339	4.350
East Pre-1971	2.804	18.971
East Post-1971	3.930	89.427

Note: The first panel of this table presents estimated results of parameters in the above equation. Results and standards errors in the parentheses are based on 25 random draws of marginal and fixed costs from previous subsections, and also 50 random draws of the simulation sample v. Mean J-stat for the west market is 2.478 and mean J-stat for the east market is 1.299. (*** p<0.01, ** p<0.05, * p<0.1). The specification of the outside benefits can be interpreted in the way that the normal distribution of outside benefits depends on the regional market and vintage group. By linear transformation the second panel of this table shows the distribution parameters for each group.

Table 2.12: Fitness of the Estimation of the Outside Benefits

	Pre-1971 Operating Units			
	West Data	West Simulated	East Data	East Simulated
Count	32	32	82	82
Capacity (MW)	159	148	279	284
Initial Year	1962	1962	1964	1964
	Post-1971 Volunteer Unit			
	West Data	West Simulated	East Data	East Simulated
Count	1	2	8	8
Capacity (MW)	54	244	635	513
Initial Year	1978	1981	1979	1978

Note: This table shows the average simulated moments and the data moments to examine the fitness of fit of the estimation results in Table 2.11.

2.5 Counterfactual Simulations

This section will evaluate how the mixed policy instruments within the voluntary cap-and-trade programs affect the regulation efficiency and unit closure when compared to a single mandatory cap-and-trade policy. I simulate the equilibrium outcomes under the counterfactual regimes listed in Table 2.13. Because the observed framework is rather complicated with overlapping and multiple policies, I change the regulatory regimes step by step to decompose the welfare changes.

Table 2.13 starts from the observed regulatory framework that pre-1971 units are mandated in the cap-and-trade program with two emission markets and post-1971 units select whether to volunteer for it. Additionally, the cap-and-trade and east intensity regulation are overlapping with each other. In regime 2, the overlapping restriction is removed to quantify its welfare consequences. Also in regime 2, the regulatory framework becomes a standard voluntary cap-and-trade framework. The regime 3 is where all the units are mandated to participate in the cap-and-trade program with the two existing regional emission markets, and the units receive allowances based on the allocation rule stated in Chapter 1. Comparing regimes 2 and 3 can let us learn whether mixed instruments under the voluntary regime dominate the single cap-and-trade instrument. In regime 4, there is only one cap-and-trade market where east and west units can freely trade emission allowances with each other. This regime is to estimate the welfare change caused by the segregation of emission market.

The welfare measures to compare the cost efficiency of these counterfactuals include the profits (including the outside benefits), emissions and an overall welfare measure that equals to the profits minus the monetary value of emission damage. The value of emission damage is set as $2640 per year per ton. It is taken from Muller (2011) that estimates the marginal damage of NO_x emissions by Texas electricity generating units using the data from approximately the same sample period as in this chapter. The value represents the marginal damage of NO_x emissions per ton to human health within a year. Although the damage of NO_x emissions mostly affects the population lives in places where they are generated, NO_x emissions can also travel long distances with the wind so that the damage is not confined to the emission source. Therefore, it is appropriate to use a constant value to measure the emission damage. I will also investigate the spatial distribution of emissions to test whether the changes in policy instruments have significant impacts in

Table 2.13: Counterfactual Policy Regimes

	Pre-1971	Post-1971	CAT Markets	Overlapping Policy	What to Evaluate
1	CAT or exit	Self-Selection	Two	Yes	Observed Framework
2	CAT or exit	Self-Selection	Two	-	No Overlapping Policies
3	CAT or exit	CAT	Two	-	Mixed vs. CAT
4	CAT or exit	CAT	One	-	Separated Markets

CAT: cap-and-trade. 'Yes' in column 'Overlapping' indicates the east intensity regulation also applies to CAT participants.

the low air quality areas in the east Texas. Finally, I compute the fraction of exit pre-1971 units to evaluate the effects of mixed policy instruments on market exit and industry structure.

During the computation of all counterfactual results, I assume that the electricity prices remain unchanged and the generating units in the cap-and-trade program will use all the allocated allowances. To compute the standard errors, I randomly draw 25 sets of parameters of the emission function, marginal and fixed costs from the estimated asymptotic distributions in the previous section and repeat the same estimation procedure for every set. These parameters are the same as those used for the estimation of the outside benefits. The appendix section gives the detailed procedure of the counterfactual computation.

2.5.1 Aggregate Outcomes

Table 2.14 summarizes aggregate welfare in each regulation regime listed in Table 2.13 and Table 2.15 summarizes the changes in welfare with the changes in regulatory regimes step by step. The results in Table 2.15 show that total emissions are very similar under the all the counterfactual regimes, but the total profits differ a lot. Relaxing the existing overlapping restrictions on the cap-and-trade units increases the aggregate welfare and profits by around $165m per year without significant impact on aggregate emissions. Changing from the voluntary cap-and-trade regime to the mandatory cap-and-trade regime increases the welfare by $176m per year, but not statistically significant. When the two regional cap-and-trade markets are merged into one, because the single market regime expands the size of participants to trade the allowances and reduce emission costs, the welfare increases further by \$199m per year with both the reduction in aggregate emissions and the increase in aggregate profits.

Table 2.14 and Table 2.15 also show that the mandatory cap-and-trade regulation reduces the exit rate of pre-1971 units by around 14%. To investigate which type of pre-1971 units change the exit choice, I compare the characteristics of the exit units under the observed policy framework with those units change the exit choices under the mandatory cap-and-trade regulation. The comparison results are presented in Table 2.16. It suggests that those units change their exit decisions under the mandatory cap-and-trade regulation are relatively larger in capacity, more fuel-efficient compared to those units still choose to exit from the market. They are more concentrated in the Houston-Galveston and Dallas Fort-Worth area and belong to the utility firms of Texas Genco, AEP Texas North and TXU. Table 2.16 implies that the mandatory cap-and-trade regulation is effective at driving those inefficient and small pre-1971 units to exit.

In terms of the aggregate cost efficiency, we can conclude that the mandatory cap-and-trade instrument dominates mixed instruments within the voluntary cap-and-trade program. This is consistent with the analysis in the conceptual framework section. There are low-cost post-1971 units staying outside of the cap-and-trade, and their non-volunteer decisions fail to fully inter-

Table 2.14: Simulation Results of Aggregate Welfare

Regime	(1) Observed Framework	(2) No Overlapping Policies	(3) Mandatory CAT Two Markets	(4) Mandatory CAT One Market
Aggregate Emissions (k tons)	206	206	204	202
	(2.88)	(3.02)	(1.04)	(1.03)
Profits ($m)	2,751	2,900	3,071	3,277
	(295)	(281)	(316)	(322)
Welfare ($m)	2,207	2,358	2,534	2,749
	(289)	(275)	(317)	(323)
Fraction of Exit Pre-1971 Units	24.1%	24.1%	9.5%	10.5%
	(0.5%)	(0.5%)	(0.4%)	(0.2%)

Note: Each counterfactual regime is defined in Table 2.13 respectively. The welfare is defined as the difference between profits and the monetary damage value of emission. The results are measured on an annual basis by using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The medians and standard errors (in the parentheses) of the 25 sets of counterfactuals are reported.

Table 2.15: The Decomposition of Changes in Aggregate Welfare

	from Observed Framework to Single Cap-and-Trade Market			
Regime Change	Total Change (4)-(1)	Removing Overlapping Policies (2)-(1)	from Voluntary to Mandatory (3)-(2)	from Two Markets to One Market (4)-(3)
Changes in Emission (k tons)	-4.6	0.1	-2.7	-1.8
	(3.4)	(0.6)	(3.4)	(1.0)
	-2.2%	0.0%	-1.3%	-1.0%
Changes in Profits ($m)	538	165	171	197
	(175)	(42)	(170)	(36)
	19.6%	6.0%	6.2%	7.2%
Changes in Welfare ($m)	549	165	176	199
	(172)	(42)	(167)	(36)
	24.9%	7.5%	8.0%	9.0%
Changes in Exit Rate	-13.6%	0	-14.6%	0.9%
	(0.6%)	(0.2%)	(0.7%)	(0.5%)

Note: The column 'Total Change' computes the changes in welfare from regime (1) to (4) in Table 2.13. The column 'Removing Overlapping Restrictions' computes the changes in welfare from regime (1) to (2). The column 'from Voluntary to Mandatory' computes the changes in welfare form regime (2) to (3). The column 'from Two Markets to One Market' computes the changes in welfare form regime (3) to (4). The results are measured on an annual basis by using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The medians and the standard errors (in the parentheses) are based of the 25 sets of counterfactuals are reported. The percentage changes are all computed relative to the median aggregate welfare measures in the observed regime (1) in Table 2.14.

Table 2.16: Characteristics of Exit Units

	Initial Service Year	Capacity (MW)	Efficiency (MWh/MMBtu)	West	East and Central	Houton Galveston	Dallas Fort-Worth
Observed Exit Units	1960	172	0.078	30%	42%	15%	12%
	Owner:	Texas Genco, AEP Texas North, TXU, San Antonio, Topaz Power Group					
Observed Exit Units	1960	194	0.082	39%	11%	22%	28%
Stay in mandatory CAT	Owner:	Texas Genco, AEP Texas North, TXU					

Note: This table summarizes the characteristics and firm ownership of observed exit units, and those observed exit units that choose to stay in the market under the mandatory cap-and-trade regulation.

nalize all the benefits of the emission abatement in the cap-and-trade program. Thus, more welfare gain could be obtained by changing into a mandatory cap-and-trade regulation. With more units in the cap-and-trade market, all the units could seek for more efficient allocation of emission abatement and production, which makes it more profitable for some pre-1971 units to stay in the market to continue operating. As a result, fewer pre-1971 units exit the market under the mandatory cap-and-trade regulation. Still, it remains a question whether the aggregate welfare gains are equally distributed. The rest of this sections will present the distributional results. There are two distributional results worth highlighting. The first is that accounting for the changes in the spatial distribution of emissions, separating the emission markets might be more desirable than the single market regime under the mandatory cap-and-trade regulation. Another is that only several large and low-cost units owned by a small number of firms gain in profits under the mandatory cap-and-trade regulation. The mixed policy framework is effective at helping small and high-cost units to obtain higher profits. Therefore, changing into the mandatory cap-and-trade regulation does not necessarily lead to Pareto improvements.

2.5.2 Distributional Results

On the distributional results, I firstly examine how the counterfacutal regimes affect the spatial distribution of emissions. Table 2.17 presents the quantity of emissions in west and east Texas under each counterfactual regime. It shows that removing the overlapping restrictions do not significantly change the total emissions of east Texas. When changing into the mandatory cap-and-trade regime, the east market has minor reductions in emissions but the total emissions in the west markets remain the same. There are two channels of effects on the emissions when changing from the voluntary cap-and-trade regime to the mandatory regime. On one hand, if the total allowance allocations to the post-1971 non-volunteer units are lower (or higher) than their emissions under the intensity standards regulation, total emissions under the mandatory cap-and-trade would decrease (or increase). On the other hand, if more (or fewer) pre-1971 units choose stay in the market under the mandatory cap-and-trade, the total emissions would increase (or decrease) since this chapter assumes that exit units' allowances are forfeit.

To compare the two channels of effects explained above, Table 2.17 also presents the distributions of emission across the vintage. Comparing the results in columns (2) and (3) we can conclude that the first channel of effect dominates the second in the east Texas. However, in the west Texas the post-1971 units increase their emissions but pre-1971 units reduce their emissions. This is caused by the cost difference between this two groups of units, which could be explained by Table 2.8. As the cost shock could represent whether each unit is the high-cost or low-cost type, Table 2.8 shows that the non-volunteer post-1971 units in the west Texas have higher costs compared to other units in the west. This implies they would have a higher demand of emission allowances and increase their emissions in the cap-and-trade. Contrarily, for non-volunteer post-1971 units in east Texas, they are relatively cost-efficient compared to the pre-1971 units in east Texas, and in cap-and-trade equilibrium they make more emission abatement to sell allowances to pre-1971 units. This leads to the east pre-1971 units to be more likely to stay in the market and increase their emissions.

Another observation from Table 2.17 is that when the two emission markets merge into one market, emissions from east units go up. It is driven by the changes in the equilibrium prices as shown in Table 2.19. When the east and west markets are segregated, the prices in the west are lower than the prices in the east. Hence, when they merge together the new price will lie in the middle, and then the west units will become sellers of emission allowances.

To further investigate how the regulatory changes affect the spatial distribution of NO_x emissions in the east Texas where the ozone attainment status is a serious concern, Figure 2.2 presents the Choropleth Maps for the percentage changes of emissions by county when changing the regulatory regimes step by step. The first one depicts the changes of emissions from the regime (1) to the regime (2) when removing the overlapping restrictions of the intensity standards regulations on cap-and-trade participants. Counties in the north east experience large reductions in emissions. Together with results Table 2.17 we can conclude that these counties are where post-1971 units are more concentrated. Counties near Austin and Houston-Galveston have a large increase in the emissions, and it is potentially caused by the fact that a lot of pre-1971 units are located in these areas and they increase their emissions when removing the overlapping restrictions. The second figure depicts the changes of

Table 2.17: The Distribution of Emissions (thousand tons)

Regime	(1) Observed Framework	(2) No Overlapping Policies	(3) Mandatory CAT Two Markets	(4) Mandatory CAT One Market
West Pre-1971	20	20	16	6
West Post-1971	38	38	42	21
West Total	58	58	58	27
East Pre-1971	42	44	50	68
East Post-1971	106	104	96	107
East Total	148	149	146	175

Note: Each counterfactual regime is defined in Table 2.13 respectively. The results are measured on an annual basis by using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The medians of the 25 sets of counterfactuals are reported.

Table 2.18: Distribution of Profits By Region-Vintage ($m)

Regime	Observed Framework	No Overlapping Policies	Mandatory CAT Two Markets	Mandatory CAT One Market
West Pre-1971	47	47	50	48
West Post-1971	156	156	117	124
East Pre-1971	1,012	1,089	1,120	1,129
East Post-1971	1,548	1,579	1,810	2,007

Note: Each counterfactual regime is defined in Table 2.13 respectively. The results are measured on an annual basis by using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The medians of the 25 sets of counterfactuals are reported.

Table 2.19: Market Clear Prices ($/ton)

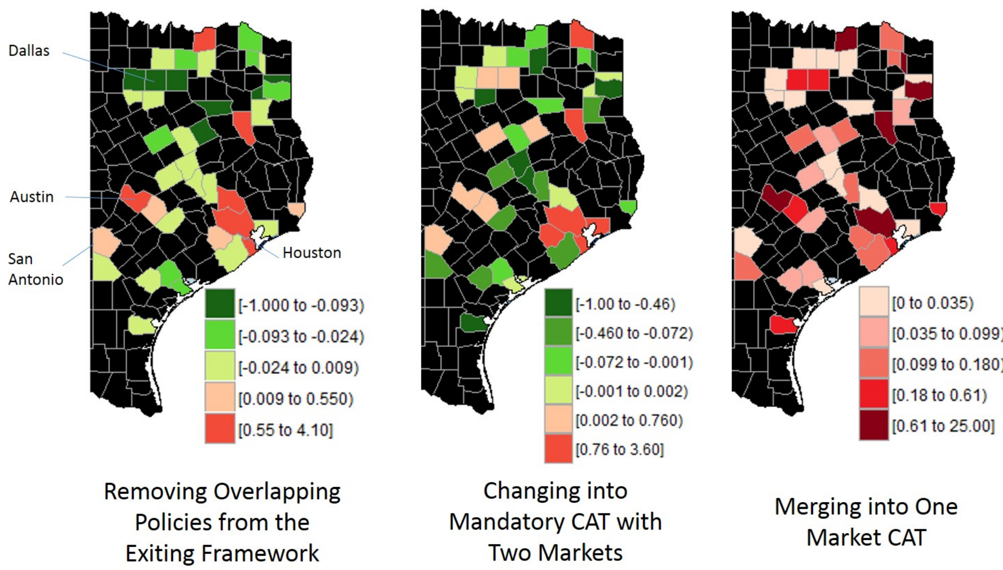
Regime	Observed Framework	No Overlapping Policies	Mandatory CAT Two Markets	Mandatory CAT One Market
West	[26, 690]	[26, 690]	[420, 690]	[966, 1246]
East	[244,1804]	[1092,2192]	[1390, 1668]	

Note: This table shows the lower and upper bounds of market-clearing prices in each emission market in the counterfactual regimes before the cap-and-trade participation decisions are made. Results are computed using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The medians of the 25 sets of counterfactuals are reported. In each emission market of each regime, there is one equilibrium price associated with one particular combination of cap-and-trade participants. Therefore, the set of equilibrium prices is not a singleton. In the last column there is only one emission market.

emissions from the regime (2) to the regime (3) when changing into the mandatory cap-and-trade regulation. Again, the majority of counties have reductions in emissions. As explained above, this reduction mainly comes from the post-1971 units. The Houston-Galveston area also experiences a huge increase in the emissions contributed by the pre-1971 units. Finally, the figure on the right shows that the one emission-market regime results in significant increases in emissions for most of the counties in east and central Texas. Considering the large size of the population in those counties, the extra emissions might cause serious health problems. This figure explains why the Texas regulator divided the cap-and-trade program into two separated emission markets.

The above results show that the intensity standards regulation in east Texas is more effective to reduce local NO_x emissions in the low air quality area Houston-Galveston compared to a cap-and-trade regulation. This difference is mainly driven by the spatial distribution of generating units that more high-

Figure 2.2: Changes in Emissions in East Texas ∗100%



Note: This figure depicts the changes of emissions from electricity generating units in the east Texas caused by the changes in the policy regime. The first figure depicts the changes of emissions when changing from regime (1) to regime (2) Table 2.13. The second figure depicts the changes of emissions when changing from regime (1) to regime (2) Table 2.13. The third figure depicts the changes of emissions when changing from regime (1) to regime (2) Table 2.13. The counties without any generating units in the sample are plotted with the black color. The results are measured on an annual basis by using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The numbers represent the medians of the ratios of the changes in emission quantity over the quantity of emissions per county in the regulatory regime (1) for the 25 sets of the counterfactuals.

cost pre-1971 units are concentrated in this area. A more general result of this section is that under the mandatory cap-and-trade regulation, the regime with separated emission markets is more effective to achieve air quality attainment in east Texas compared to the one-market regime. The policy implication is that for pollutants like NO_x with serious local damage, a uniform cap-and-trade policy might cause an undesirable spatial distribution of emission damage and exacerbate the problem of ozone non-attainment.

I also examine whether the changes in profits for individual units align with the changes in aggregate profits, and which group of units benefits most under the mandatory cap-and-trade regulation. Table 2.18 shows the distributions of profits under different regimes. The profits of the west units are always smaller than the east units. One reason is that more units are concentrated in the east and another reason is that the west units typically have higher cost and lower efficiency as shown in Table 2.8. Removing the overlapping restrictions benefits east pre-1971 units, as the intensity standards impose high penalties for these units. Changing into the mandatory cap-and-trade regulation benefits the majority of units except for post-1971 units located in the west Texas, and post-1971 units in the east Texas enjoy most of the profits gains. As Table 2.8 indicates, post-1971 units in the west Texas have the highest abatement costs among all the units, and they are unlikely to volunteer for the cap-and-trade regime under the voluntary regime. This implies that the mandatory cap-and-trade regulation is more costly for them. When merging the west and east emission markets into one market, east units and west post-1971 units experience increase in profits, and again post-1971 units in the east Texas enjoy most of the profits gains. For the east units, such gains come from the reduction of emission price as Table 2.19 shows. For the west pre-1971 units, the changes in emission price increase the abatement costs for them and reduce their profits.

As the same utility firm usually own units of different vintages and locations, investigating the distributional effects by utility firms can also help to understand the welfare effects of policy instruments. Figure 2.3 depicts the percentage changes in profits of each utility firm when changing the regulatory regimes. The first figure shows the changes of profits when removing the overlapping policies from regime(1) to regime (2). The second figure depicts the changes in profits when changing from the regime (2) to mandatory cap-and-trade regime. The third figure depicts the changes when merging into one emission market under the cap-and-trade regulation.

Figure 2.3 suggests that removing the overlapping restrictions or merging the emission markets have limited impacts on the profits for the majority of utility firms. Several firms gain in profits from regime (1) to voluntary cap-and-trade regime (2). An example is Entergy Gulf States Inc that owns units located around the Houston-Galveston area. When units in the cap-and-trade market no longer have to pay for the intensity standards violation, they increase their emission level and pay lower abatement costs.

The most important message of Figure 2.3 is that when changing into mandatory cap-and-trade regulation (represented by the middle figure), only a few firms gain in profits while others incur huge losses. These differential changes in profits could be largely explained by the characteristics of generating units owned by the firms. The firms are divided into four groups and the characteristics of their units are summarized in Table 2.20. The first group of firms like Texas Genco and Texas Municipal Power Agency remain to be the winners when changing into the mandatory cap-and-trade regime. These winner firms own relatively newer, more efficient and less costly units compared to the sample of units. Although the mandatory cap-and-trade might impose stringent quantity targets, these units still have the cost advantage over other units to compete in the emission market to ensure increases in profits. The second group of firms such as Greenville Electric Utility Sys and Sempra Energy Resources only own post-1971 units. Compared to the sample average, their units are smaller in size, less-efficient and has relatively higher abatement costs. Consequently, they are unwilling to volunteer for the cap-and-trade. When changing into the mandatory cap-and-trade regime, they are mandated to pay higher costs for emission abatement and suffer from profit loss. The third group of units, including Austin City Energy, TXU and Lubbock City Energy, owns the volunteer post-1971 units. Their post-1971 units are relatively more efficient and less costly so that by volunteering for the cap-and-trade they help the pre-1971 units reduce abatement costs. Under the mandatory cap-and-trade regime, with more units in the emission markets, they no longer remain as the mostly profitable group and they also suffer from profit loss. The last group of firms has a large number of both pre-1971 and post-1971 units, but their units do not have any cost advantages compared with others. When changing into the mandatory cap-and-trade regime, the stringent emission quantity targets induce their post-1971 units to incur higher costs to make emission abatement compared to the intensity standards[5].

Figure 2.3 also suggests that when merging the two emission markets into one market under the mandatory cap-and-trade regulation, there are also unequal distributional effects despite the increase in aggregate profits. For firms in the west Texas like El Paso Electric Co, the increase in emission price induces them to pay for higher costs and obtain lower profits. For firms like Sweeny Cogeneration LP and Texas Municipal Power Agency who are the allowance sellers in the east Texas under the two-market regime, with the merge of emission markets and the decline in the emission price they incur losses in allowance revenue. However, some firms like Austin City Energy and Southwestern Electric Power Co that own units in east Texas gain higher profits, as the emission prices decrease the abatement costs lower for them. Finally, If we combine the last two graphs in Figure 2.3, we can conclude that even with a single-market mandatory cap-and-trade regulation, a majority of firms and generating units still suffer from profit loss compared to the mixed policy framework.

[5]Although Firms like Lower Colorado River Authority and Garland City experience increases in profits when changing into the mandatory cap-and-trade regime, their original profits under the voluntary cap-and-trade regime were estimated to be negative, which implies that they are the least profitable firms. Changing into mandatory cap-and-trade regime leads these firms to cut production to prevent further loss, but they are not as profitable as the first group of firms. Therefore, these firms are classified into the fourth

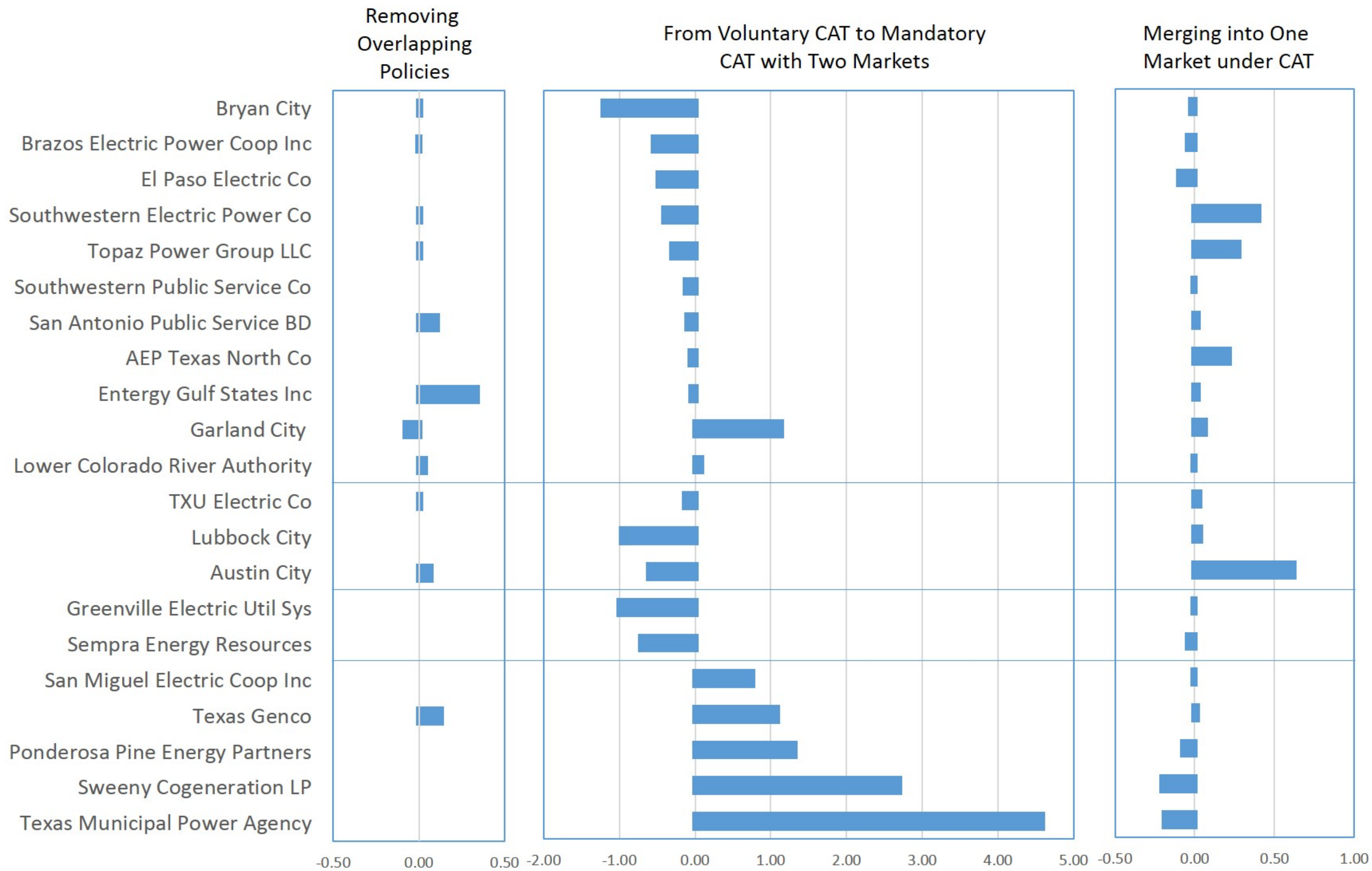
Table 2.20: Characteristics of Generating Units by Group of Firms

	Initial Service Year	Capacity (MW)	Efficiency (MWh/MMBtu)	Cost Shock ($/MMBtu)
All Pre-1971	1963	230	0.086	0.049
Group1 Pre-1971	1963	340	0.091	-0.039
Group3 Pre-1971	1962	280	0.084	0.015
Group4 Pre-1971	1963	164	0.085	0.095

	Initial Service Year	Capacity (MW)	Efficiency (MWh/MMBtu)	Cost Shock ($/MMBtu)
All Post-1971	1981	416	0.090	-0.109
Group1 Post-1971	1985	464	0.098	-0.216
Group2 Post-1971	1982	104	0.083	0.139
Group3 Post-1971	1978	570	0.091	-0.145
Group4 Post-1971	1979	384	0.090	-0.073

Note: This table shows characteristics of units of each group of firms in Figure 2.3. The first group includes Texas Municipal Power Agency, Texas Genco, Sweeny Cogeneration LP, Ponderosa Pine Energy Partners and San Miguel Electric Coop Inc. The second group includes Greenville Electric Utility System and Sempra Energy Resources. The third group includes Austin City, TXU Electric Co and Lubbock City. The fourth group includes all the rest of the firms.

Figure 2.3: Percentage Changes in Profits by Utility Firms



Note: This figure shows the percentage changes in profits of each utility firm when changing from the benchmark to the voluntary cap-and-trade regime and when changing from the voluntary cap-and-trade regime to the mandatory cap-and-trade regime. The results are measured on an annual basis by using 2004 data and 25 random draws from the estimated asymptotic distribution of the parameters of the emission function, marginal and fixed costs from previous subsections. The medians of the 25 sets of counterfactuals are reported.

This subsection reveals that when changing into mandatory cap-and-trade regulation, the individual profits do not align with the aggregate increase in profits. The drop in profits under the mandatory cap-and-trade regime for a large number of firms explains why we observe such few volunteers in the Texas cap-and-trade program. It also implies that even if the Texas regulator intended to adopt a mandatory cap-and-trade regulation, the regulation would not be supported by the majority of utility firms. It suggests that to make the mandatory cap-and-trade regulation a Pareto improvement over the existing mixed policy framework, re-distributional policies are necessary to align the changes of individual profits with the aggregate gains.

2.6 Concluding Remarks

This chapter studies the effects of mixed policy instruments in the voluntary cap-and-trade programs on social welfare and unit closure using data from the Texas electricity generation industry. The generating units built before 1971 were mandated to participate in a cap-and-trade program to reduce their NO_x emissions, while units built after 1971 could choose between volunteering for this cap-and-trade program or not. In the meantime, units in east Texas are regulated by a separate intensity standards regulation. Given this background, this chapter builds a structural model of cap-and-trade participation and abatement choices of generating units. In the model, units firstly

group.

select their preferred policy instrument and then choose the abatement technology and production. Based on the variations from the observed choices, this chapter estimates the abatement cost parameters of generating units and use the parameters to conduct counterfactual analyses.

Counterfactual analyses indicate that a single cap-and-trade instrument outperforms the mixed instruments within the voluntary program in terms of aggregate cost efficiency, and also reduces the market exit rate. The result is driven by the mechanism that in the market-based cap-and-trade regulation, units can efficiently allocate the production and emission abatement through the emission market transaction, but the individual self-selection decisions fail to internalize such overall benefits. However, the mandatory cap-and-trade instrument mainly benefits large and low-cost units and induce other units to have lower profits. This chapter also discovers that within the mandatory cap-and-trade regime, dividing the cap-and-trade program into two segregated markets could reduce the emissions generated from a majority of counties in east Texas while regulating by the single emission market has the opposite effect. Some of these counties failed to meet the federal air quality standards. Hence, a cap-and-trade program with separated emission markets is more effective for them to achieve the attainment status.

This chapter contributes to the literature on the choices of environmental policy instruments. Two potential policy implications could be drawn from the analysis. Firstly, this chapter indicates that a uniform policy instrument dominates the vintage-differentiated policy instruments. Vintage-differentiated environmental regulations have a long history in the world, but this chapter provides new evidence of the welfare consequences of this type of differentiated regulations. Moreover, to remove the past grandfathering provisions, regulating all the units by cap-and-trade will outperform the creation of a separate cap-and-trade market for the previously grandfathered units. It will also reduce the likelihood of exit for those older units.

Another policy implication of this chapter relates to the continuing debates on policy instruments choice. This chapter discovers that regulating every unit by the cap-and-trade instrument achieves higher aggregate efficiency compared to the mixture of instruments. There are also important distributional effects on the regulated firms. The regime improves the aggregate welfare might cause a large number of firms to incur profit loss, and that might cause the regulated firms to fight against the new mandatory regulations. Re-distributional policies are necessary to align the aggregate welfare with individual welfare to eliminate such conflicts. Finally, although the regulations studied in this chapter are enforced at the state level, the implications could also be applied to regulations at larger regional levels such as the U.S. and the European Union.

This chapter studies the differentiated emission regulations in the electricity generation industry. The energy industries are heavily regulated in various dimensions. Future work could investigate further into the impacts of interactions between differentiated environmental regulations and other price or non-price regulations, and also how the environmental regulations affect the vertical and horizontal structures of the energy industries.

Chapter 3

The Impacts of Congestion Pricing on Emission Externalities and Fuel Efficiency: Evidence from the ERCOT

3.1 Introduction

Transmission congestion is a critical issue in the electricity market. Whenever it happens, the low-cost generator is unable to provide energy to the consumers and the total cost of the electricity system is not minimized. There are two common ways to incorporate congestion cost into the wholesale electricity price: zonal and nodal market structures. Under the zonal market structure, the market is divided into separated loading zones and the electricity price is the marginal cost of production of the entire zone. The zonal market structure can use different zonal prices to deal with inter-zonal congestions but not intra-zonal congestions. Contrarily, the nodal market structure is able to deal with all types of congestions with node-specific prices. Nowadays, the nodal market structure is prevalent among the electricity markets under the argument that it is more effective at providing price signals and minimizing the total cost of the market. However, with the presence of transmission network constraints and outages caused by the construction of the new transmission lines, the nodal market structure might not be as efficient as the above argument predicts.

This chapter addresses the question that whether the form of transmission congestion pricing affects the emission externalities and fuel efficiency of electricity generation using evidence from the ERCOT (Electric Reliability Council of Texas). The ERCOT is one of the most competitive electricity markets in the North America. Starting in the year 2002, ERCOT adopted a zonal market structure to price transmission congestion because the regulator believed that inter-zonal congestions would happen more frequently than intra-zonal congestions, and that a zonal market design would be more effective to manage inter-zonal congestions. With time went by, the regulator realized their projections on the type of congestion were wrong and that the zonal market structure led to tremendous cost for the market participants. Then, in the year 2010, the ERCOT changed into a nodal market structure.

I use a difference-in-difference approach to empirically investigate the effect of the regulatory changes on efficiency outcomes. As explained in Chapter 1, without congestion the two market structures will lead to the same price and dispatch order so that only congestion areas are affected by the choice of market structure. Therefore, I classify those power plants that located in counties with serious congestion as the treatment group of the nodal market structure. I collect daily and monthly emission and fuel usage data of power plants located in ERCOT during the summer season in 2009 and 2011, which includes periods before and after the regulatory change. I pick the summer time for comparison because transmission congestion is more likely to occur during this season compared to other seasons in Texas. I find that the nodal market structure has heterogeneous impacts on areas with different causes of transmission congestions. For counties located along the path to transferring wind generation from west to east Texas, the nodal pricing leads to increases in emission intensities of fossil-fired power plants, although the total increase in emission cost is not economically significant. Contrarily, the nodal pricing increases the fuel efficiency by 2-9.6% for power plants located around big cities with excess demand during summer time, and the estimated fuel cost saving is around $154.8m.

The empirical question is motivated by the literature on the economic regulations in the electricity sector (Fabrizio, Rose and Wolfram (2007); Davis and Wolfram (2012); Chan, Iange and Li (2013); Abito (2014); Hausman (2014); Cicala (2015); Jha (2015); Lim and Yurukoglu (2015)). The literature follows the argument in Laffont and Tirole (1993) that price regulations discourage cost-minimizing behaviors, and that deregulation brings more efficiency gains. Hausman and Muehlenbachs (2016) applies similar arguments to examine the emission externalities in the natural gas market. The contribution of this chapter is that I use the emission intensities and fuel efficiency as the measure of cost-minimizing efforts, and that I compare such efficiency outcomes across different congestion pricing structures in a competitive wholesale market.

This chapter also contributes to the literature on environmental and price regulations in the electricity industry. Previous research focuses on the effects of economic regulations on the cost-effectiveness of emission trading programs. For example, Fowlie (2010) addresses the impacts of electricity restructuring on abatement technology investment and the performance of the emission trading regulation using data from the Ozone Transport Commission NOx Budget Program. Limpaitoon, Chen, and Chen (2010) uses a simulation study to investigate the effects of the emission trading on the congested electricity market. The most related paper is Zhang(2016), which uses 2010-2011 data from ERCOT to estimate the cost savings of the new market structure. However, there are several major differences between my paper and his paper. First of all, I use a difference-in-difference approach by comparing the cost-minimizing efforts in congestion areas versus non-congestion areas before and after the regulatory change, but his paper is focusing on the impacts on production decisions. Secondly, I use novel data sources to control the impacts of changes in transmission network and different types of causes of transmission congestion. Moreover, my results show that even though the nodal market brings clear price signals to buyers and sellers in the wholesale market, the network constraint in west Texas could become a big obstacle for the nodal market structure to obtain further efficiency gains. This provides important policy implications for the ERCOT to improve their planning for the future.

This chapter contributes to the literature on transmission network constraint and the competitiveness of electricity markets as well (Hyunsook (2002); Joskow and Tirole (2000); Neuhoff et.al (2005); Cho and Kim (2007)). Joskow and Tirole (2000) gives the theoretical foundation of the network effects and the optimal design of transmission rights market. Other papers concentrate on the effects of network constraint on the market power by examining the bidding behaviors in the wholesale market, and the empirical papers mainly use data from the Europe or the California market for quantitative analysis. This chapter adds to the literature by examining the network effects on cost-reducing efforts, measured by emission intensity and fuel efficiency with new evidence from the ERCOT. My results show that in areas with congestion caused by excess renewable sources and insufficient transmission capacity, the nodal market structure would lead to increasing emissions from fossil-fired sources.

As one of the most successful electricity market in the world, ERCOT has always been the focus of empirical research to improve our understanding of electricity market structure. There are also a lot of related papers studying the pricing of the ERCOT market. For instance, Hortasu, Madanizadeh and Puller (2015) studies the bidding behaviors in the wholesale market; Roderick (2013) studies the Ramsey problem in the transmission sector; Hortacsu and Puller (2008) studies the impact of deregulation of retail market on the retail prices; Baldick et.al (2014) quantifies the impact of the nodal market design on the wholesale prices; and Cullen (2014) uses the ERCOT's data to study the dynamic production decision of fossil fuel power plants in response to the carbon regulations and wind generation. This chapter adds to the literature by examining how congestion pricing in the ERCOT affect the power plant operation efficiency.

This chapter is organized as the following. The next section introduces the data, and the following section specifies the identification strategy. Then, this chapter gives and interprets the estimation results. The last section concludes.

3.2 Data

3.2.1 Data Source

The data used in this chapter are obtained from the Energy Information Administration's forms (EIA-767/860, 923, 423/920/926) and the EPA Air Market Program database. The EIA data provides me with monthly generation and fuel use data for the electric generating facilities in the ERCOT with a total generator nameplate capacity of 1 or more megawatts. The EPA database includes daily emission and generation information for fossil-fueled generating facilities in the ERCOT. As mentioned earlier, the nodal market structure leads to different prices compared to the zonal market structure only when transmission congestion occurs. In this chapter, I use county-level congestion information taken from the *Report on Existing and Potential Electric System Constraints and Needs* provided by the ERCOT for the years 2009-2011. These reports include the location of significant congestion happening in each year and the primary cause of the congestion. They also include the location of newly completed or upgraded transmission lines in the ERCOT in each month.

The sample period is from the June to the August of the years 2009 and 2011, which consists of the summer season before and after the nodal market transition. I choose the summer season because the electricity demand is the highest during the summer in Texas so that it is more likely to observed congestion during such period. Without transmission congestion, even under the nodal market structure every node's price will be the same within the same zone so that the two pricing strategies will have the same impacts on the grid. During the sample period, all the power plants in Texas are subject to the federal emission regulation, the Clean Air Interstate Rule (CAIR). Almost all of them have finished upgrading their emission control technologies before the year 2009 so that using this period of data does not require modeling the effects of regulations on the choice of emission control technologies, and that emission intensity[1] is a good indicator of the emission externalities and emission reduction efforts.

I also collect weather data to use as the proxy for electricity demand and wind generation potential during the sample period. The weather data includes daily average temperature and wind speed, and they are taken from the website *http://www.almanac.com/weather*. For each power plant in the EPA sample, I use the weather data from the nearest city weather station.

3.2.2 Summary Statistics

Table 3.1 and Table 3.2 give the summary statistics of the two samples used in this chapter. Table 3.1 table shows the summary statistics of data

[1]The ratio between emission and production or heat input.

Table 3.1: EPA Sample Summary Statistics (Daily-Plant Obs)

Variables	Obs	Mean	Std. Dev.	Min	Max
Generation (MWh)	16,173	11,462.1	12,276	0	74,929.1
Heat input (MMBtu)	16,188	105,907.4	117,195.6	0	752,979
NOx (tons)	16,188	4.697	9.000	0	57.886
Daily mean temperature	17,235	86.490	4.107	66.9	105.8
Wind speed (Mph)	17,197	8.312	3.422	0	26.24
Congestion Indicator	17,235	0.594	0.491	0	1
Border_Congestion Indicator	17,235	0.048	0.214	0	1
Local_Congestion Indicator	17,235	0.546	0.497	0	1
Wind_Congestion Indicator	17,235	0.037	0.189	0	1
Load_Congestion Indicator	17,235	0.316	0.465	0	1
Outage_Congestion Indicator	17,235	0.470	0.499	0	1
New Line Indicatior	17,235	0.375	0.484	0	1
# Unit Per Plant	17,235	3.096	2.030	1	14

Note: This table shows the summary statistics of data using the EPA sample for all the fossil-fueled power plants in the ERCOT. This sample includes daily plant-level observations during June to August of the years 2009 and 2011. The variable 'Congestion Indicator' equals to 1 if the power plant is located in counties with transmission congestion in the year 2011, and equals to 0 otherwise. The next two variables indicate whether the congestion occurs at the border of loading zones or within a loading zone. The next three variables equal to 1 if the congestion in the year 2011 is due to the transmission of wind generation, excess demand or outage caused by the construction of new lines respectively. The 'New Line Indicator' equal to 1 if the power plant is located in counties with new transmission lines in the year 2011 emission intensities with the year 2009. The last variable measures the number of generators in a single power plant to measure the size of that power plant.

Table 3.2: EIA Sample Summary Statistics (Month-Plant Obs)

Variable	Obs	Mean	Std. Dev.	Min	Max
Generation (MWh)	1,692	142,625.4	283,301.8	-384	2,041,036
Heat input (MMBtu)	1,692	1,429,162	2,891,485	0	21,479,258
Congestion Indicator	1,692	0.489	0.500	0	1
Border_Congestion Indicator	1,692	0.027	0.161	0	1
Local_Congestion Indicator	1,692	0.462	0.498	0	1
Wind_Congestion Indicator	1,692	0.108	0.316	0	1
Load_Congestion Indicatior	1,692	0.265	0.441	0	1
Outage_Congestion Indicator	1,692	0.319	0.466	0	1
New Line	1,692	0.346	0.476	0	1
Capacity (MW)	1,692	770.679	2,194.789	1.1	32,067.2

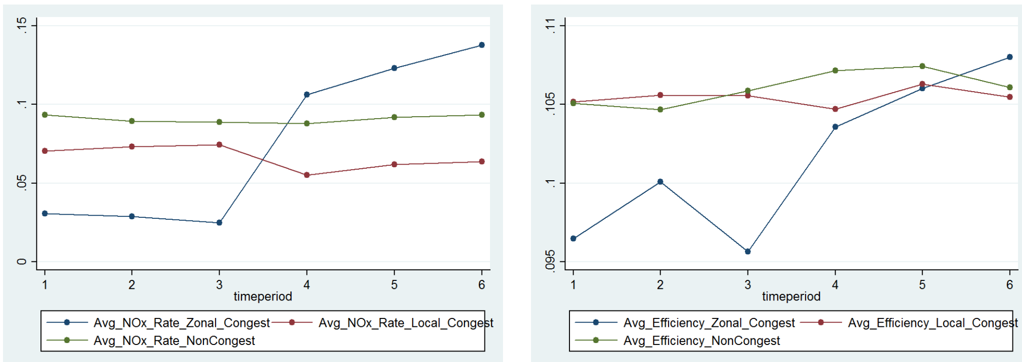
Note: This table shows the summary statistics of data using the EIA sample for all the power plants with a capacity larger than 1MW in the ERCOT. This sample includes monthly plant-level observations during June to August of the years 2009 and 2011. The variable 'Congestion Indicator' equals to 1 if the power plant is located in counties with transmission congestion in the year 2011, and equals to 0 otherwise. The next two variables indicate whether the congestion occurs at the border of loading zones or within a loading zone. The next three variables equal to 1 if the congestion in the year 2011 is due to the transmission of wind generation, excess demand or outage caused by the construction of new lines respectively. The 'New Line Indicator' equal to 1 if the power plant is located in counties with new transmission lines in the year 2011 emission intensities with the year 2009. The last variable is the total nameplate capacity of the power plant.

using the EPA sample for all the fossil-fueled power plants in the ERCOT. This sample includes daily plant-level observations during June to August of the years 2009 and 2011. Table 3.2 shows the summary statistics of data using the EIA sample for all the power plants with a capacity larger than 1MW in the ERCOT. This sample includes monthly plant-level observations during June to August of the years 2009 and 2011. Comparing these tables show that power plants in the EPA sample have larger generation and use more fuel input during the summer time.

In the ERCOT market, the transmission congestion could happen at the border of the loading zones or within the loading zones. In Table 3.1 and Table 3.2, the variables *Border Congestion Indicator* and *Local Congestion Indicator* represents whether the location of congestion is at the border. As these tables show, most of the congestion happens within the loading zone (local congestion). The congestion also happens for various reasons, such as outages caused by the construction of new lines, excess demand of load, and the inability to transfer wind generation from west to east Texas given limited capacity. In Table 3.1 and Table 3.2, the variables *Outage Congestion Indicator*, *Load Congestion Indicator*, and *Wind Congestion Indicator* represent the above causes respectively. For each county, there could be multiple causes of transmission congestion. The summary statistics indicate that the excess demand and construction outages are the major causes of congestion in the ERCOT.

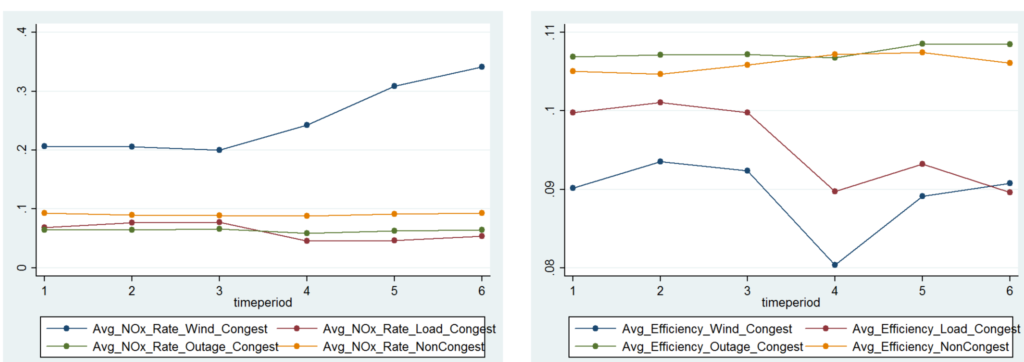
I also compare the congest counties with the non-congest counties in

Figure 3.1: Compare Congest Counties vs. Non-Congest Counties I



Note: Graph on the left depicts the average NO_x emission intensities by each group in each month. Graph on the left depicts the average fuel efficiency by each group in each month.

Figure 3.2: Compare Congest Counties vs. Non-Congest Counties II



Note: Graph on the left depicts the average NO_x emission intensities by each group in each month. Graph on the left depicts the average fuel efficiency by each group in each month.

the sample in Figure 3.1 to Figure 3.2. Before the regulatory change, the NO_x emission intensities and fuel efficiency is relatively stable. After the regulatory changes, there are heterogeneous trends of changes for different groups of power plants. Power plants in non-congest counties have higher emission intensities. After the regulatory change, the average NO_x intensity increases for power plants located in congestion counties with zonal congestion or excess wind capacity. The average fuel efficiency of power plants in zonal congest counties increases after the regulatory change, but decreases in counties with congestion caused by excess demand or excess wind capacity. These results indicate that the pricing structure is associated with heterogeneous degrees of cost-minimizing efforts by power plants located in different congestion areas. The next sections will use regression analysis to identify the impacts of the pricing structure.

3.3 Empirical Strategy

As mentioned in Chapter 1, the main difference between zonal and nodal market design in the ERCOT is that the dispatch order is on a resource-specific basis in the nodal market. It implies that prices are heavily dependent on the resource-specific marginal cost and transmission capacity. Therefore, the hypothesis of this chapter is that under the nodal market structure, the power plant managers will pay more efforts to improve generation fuel efficiency and reduce emission intensity in order to reduce their own marginal costs and increase profitability.

I use a plant-level fixed-effect model to test the hypothesis. The primary specification is:

$$Y_{ict} = \alpha_i + \theta \mathbf{1}\left(c \in congestion\right) \mathbf{1}\left(t \in 2011\right) + X_{ict}\beta + W_{ct}\gamma + Z_t\sigma + \varepsilon_{ict} \quad (3.1)$$

where i is the power plant, c is the county where power plant i is located, and t is the time period. The dependent variable, Y_{ict} represents the emission intensities of NO_x and the fuel efficiency of generator i in county c at time period t. Since the majority of fossil-fueled generating units in Texas uses natural gas, NO_x is the main type of pollutant in this market. I did not use the quantity of emission to be the dependent variable to avoid the effect of time-varying demand. Most of the generating units already retrofit their boilers before the start of the CAIR to install emission control devices. Therefore, during the sample period, the emission intensities and fuel efficiency are good indicators of unobserved emission reduction and cost reduction efforts.

For the variables on the right-hand side, X includes power plant-level variables like the number of generating units to control for the impacts of entry and exit. W includes county-level variables affecting the demand and transmission congestion of electricity, which consists of average temperature, wind speed for west Texas counties, and the dummy variable indicating whether the county has new transmission lines in the year 2011. Z includes the time fixed effects, such as indicators of the year, month or weekday. The coefficient of interest, θ, is on the interaction between an indicator for the nodal pricing period and an indicator for whether transmission congestion happens in each specific county. This coefficient thus represents whether the fuel and emission efficiency improves after the regulatory change for power plants located in congested counties compared to non-congested counties.

To distinguish between the zonal and local congestion locations, I also introduce the following specification:

$$\begin{aligned} Y_{ict} = \alpha_i + \theta_1 \mathbf{1}\left(c \in Zonal\_Congestion\right) \mathbf{1}\left(t \in 2011\right) + \theta_2 \mathbf{1}\left(c \in Local\_Congestion\right) \\ \mathbf{1}\left(t \in 2011\right) + X_{ict}\beta + W_{ct}\gamma + Z_t\sigma + \varepsilon_{ict} \end{aligned} \tag{3.2}$$

To identify whether different causes of congestion has heterogeneous impacts on the energy efficiency and emission outcomes, I also include the following regression where the treatment group is divided into subgroups based on the type of causes:

$$\begin{aligned} Y_{ict} = \alpha_i + \theta_1 \mathbf{1}\left(c \in Outage\right) \mathbf{1}\left(t \in 2011\right) + \theta_2 \mathbf{1}\left(c \in Load\right) \mathbf{1}\left(t \in 2011\right) + \\ \theta_3 \mathbf{1}\left(c \in Wind\right) \mathbf{1}\left(t \in 2011\right) + X_{ict}\beta + W_{ct}\gamma + Z_t\sigma + \varepsilon_{ict} \end{aligned} \tag{3.3}$$

3.4 Results

Table 3.3 to Table 3.5 show the regression results of Equation 3.1 and Equation 3.2. The standard errors are clustered at the county level. The nodal congestion pricing does not have significant impacts on the NO_x emission intensities. If we divide the congestion counties into those located on the borders of loading zones and those located within each loading zone, we can find that a larger proportion of NO_x emission increase comes from the loading zone borders. If we divide the congestion counties into different groups, we can find that power plants located in the counties with congestion caused by transferring wind generation increase their NO_x emission intensity. In these counties, the insufficient transmission capacity prevents the injection of low-cost and clean wind generation and leads the fossil-fired power plants to engage in inefficient fuel and emission management to gain profits. This effect shows that the network constraint may prevent the nodal market structure to gain efficiency. However, as shown in the previous sections, the total number of power plants located in this type of counties only accounts for a small proportion of all the power plants in the ERCOT. Therefore, the increase in their emission intensities will not generate huge impacts on the total emissions from the ERCOT.

Regarding the effects on fuel efficiency, the EPA sample shows no significant impacts, but the EIA sample shows that the nodal pricing in counties with excess demand leads to power plants in these areas improve their fuel efficiency by 9.5%. When excess demand causes congestion, the nodal market structure gives better price signals to buyers and power plants to inform them about the market condition, and power plants have to reduce their own marginal costs to be able to supply electricity. However, for power plants located in other congested counties with network constraints like construction

Table 3.3: EIA Sample Results of NO_x Emission Rate

VARIABLES	(1)	(2)	(3)
	log(NO_x emission intensity)		
Congestion	0.117		
	(0.113)		
Zonal_Congestion		0.269*	
		(0.145)	
Local_Congestion		0.103	
		(0.113)	
Wind_Congestion			0.281**
			(0.106)
Load_Congestion			-0.011
			(0.098)
Outage_Congestion			0.068
			(0.091)
New Line	-0.078	0.011	-0.059
	(0.098)	(0.025)	(0.095)
#Unit Per Plant	-0.078	-0.045	-0.041
	(0.098)	(0.072)	(0.075)
Normalized Temperature	0.040***	0.039***	0.035**
	(0.013)	(0.013)	(0.013)
Wind Speed in West Texas	0.015	0.015	0.009
	(0.012)	(0.012)	(0.011)
Plant Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Observations	16,164	16,164	16,164
R-squared	0.897	0.897	0.897

Note: This table shows the regression results of NO_x emission intensity using the EPA sample. The daily temperature is normalized based on the sample distribution. The variable 'Wind Speed in West Texas' measures the daily wind speed in the west Texas counties, and equals to zero if the county is located in the east Texas. Time fixed effects include indicators of year, month and Monday to Saturday. Standard errors clustered at the county level in parentheses(*** p<0.01, ** p<0.05, * p<0.1). NO_x emission intensity is defined as the ratio between NO_x and heat input.

Table 3.4: EPA Sample Results of Fuel Efficiency

	(1)	(2)	(3)
VARIABLES	log(fuel efficiency)		
Congestion	-0.009		
	(0.021)		
Zonal_Congestion		-0.004	
		(0.036)	
Local_Congestion		-0.010	
		(0.021)	
Wind_Congestion			-0.030
			(0.027)
Load_Congestion			0.014
			(0.014)
Outage_Congestion			-0.015
			(0.019)
New Line	-0.016	0.011	-0.019
	(0.025)	(0.025)	(0.015)
#Unit Per Plant	-0.017	-0.016	-0.018
	(0.015)	(0.025)	(0.025)
Normalized Temperature	0.008**	0.008**	0.008**
	(0.004)	(0.004)	(0.004)
Wind Speed in West Texas	-0.004***	-0.004***	-0.004**
	(0.001)	(0.001)	(0.002)
Plant Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Observations	16,087	16,087	16,087
R-squared	0.718	0.718	0.719

Note: This table shows the regression results of fuel efficiency using the EPA sample. The daily temperature is normalized based on the sample distribution. The variable 'Wind Speed in West Texas' measures the daily wind speed in the west Texas counties, and equals to zero if the county is located in the east Texas. Time fixed effects include indicators of year, month and Monday to Saturday. Standard errors clustered at the county level in parentheses(*** p<0.01, ** p<0.05, * p<0.1). Fuel efficiency is defined as the ratio between generation and heat input.

Table 3.5: EIA Sample Results of Fuel Efficiency

VARIABLES	(1)	(2)	(3)
	log(fuel efficiency)		
Congestion	-0.017		
	(0.042)		
Zonal_Congestion		-0.042	
		(0.041)	
Local_Congestion		-0.016	
		(0.043)	
Wind_Congestion			-0.090*
			(0.049)
Load_Congestion			0.095**
			(0.045)
Outage_Congestion			-0.073*
			(0.043)
New Line	0.031	0.031	0.017
	(0.036)	(0.036)	(0.034)
Capacity	0.006	0.005	0.012*
	(0.006)	(0.006)	(0.007)
Plant Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Observations	1,606	1,606	1,606
R-squared	0.926	0.926	0.928

Note: This table shows the regression results of Equation 3.1 to Equation 3.3 using the EIA sample. Time fixed effects include indicators of year and month. Standard errors clustered at the county level in parentheses(*** p<0.01, ** p<0.05, * p<0.1).

outages or insufficient capacity, their fuel efficiency decreases by 7 to 9%. It is because in such areas the nodal market structure fails to deploy low-cost generations to supply electricity so that power plants are shirking in improving their efficiency. Considering the fact that over 50% of total congestion is caused by excess demand, the net effect shown by Table 3.5 is that the overall fuel efficient of the ERCOT improves under the nodal market structure. Using the average fuel cost in the ERCOT and the total heat input usage in summer 2011, my estimation of the total fuel cost saving is around $154.8m.

The above results show that regulatory change to the nodal market structure has impacts on the cost-minimizing efforts to improve fuel efficiency within a single power plant. However, does nodal pricing changes the dispatch order so that more efficient power plants produce more relative to others? In Table 3.6, I use the EIA sample to test whether the nodal market structure affects the production decisions of power plant located in specific congestion areas. The dependent variables are the log of the ratio of the plant-month level generation (or heat input) to the total monthly generation (or heat input) of the ERCOT. The results show that generally power plants in the congestion areas produce the same portion of generation compared to power plants in other regions, so that the form of congestion pricing does not have significant impacts on how much to produce in the summer time.

Table 3.6: Production Decisions

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES		Generation			Heat Input	
Congestion	-0.119			-0.073		
	(0.124)			(0.124)		
Zonal_Congestion		-0.234*			-0.162	
		(0.132)			(0.137)	
Local_Congestion		-0.113			-0.069	
		(0.125)			(0.124)	
Wind_Congestion			-0.065			0.045
			(0.150)			(0.121)
Load_Congestion			0.103			0.018
			(0.111)			(0.085)
Outage_Congestion			-0.194			-0.102
			(0.125)			(0.103)
New Line	0.169	0.171	0.134	0.113	0.114	0.093
	(0.148)	(0.148)	(0.131)	(0.144)	(0.144)	(0.119)
Capacity	0.067	0.064	0.079	0.063	0.061	0.069
	(0.046)	(0.045)	(0.050)	(0.045)	(0.044)	(0.046)
Plant Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,606	1,606	1,606	1,612	1,612	1,612
R-squared	0.978	0.978	0.978	0.981	0.981	0.981

Note: This table shows the regression results of Equation 3.1 to Equation 3.3 using the EIA sample. The dependent variables are the log of the ratio of the plant-month level generation (or heat input) to the total month level generation (or heat input) of ERCOT. Time fixed effects include indicators of year and month. Standard errors clustered at the county level in parentheses(*** p<0.01, ** p<0.05, * p<0.1).

3.5 Concluding Remarks

The transition from a zonal market structure to a nodal market structure by the ERCOT in the year 2010 provides a perfect setting to study the effect of transmission congestion price regulation on emission reduction and fuel efficiency improvement in the wholesale market. Using a difference-in-difference approach, I compare the impacts of different pricing structures on congested areas versus non-congested areas. For congested areas with excess load (demand), the nodal pricing structure encourages the power plant managers to reduce costs by lowering their emission intensity and improving fuel efficiency, leading to huge savings in fuel costs. This finding is consistent with the theoretical hypothesis mentioned in Chapter 1 that nodal pricing gives better price signals to help power plant managers to minimize operating costs. However, for congested areas with limited capacity to transfer wind generation from west to east Texas, the fossil-fueled power plants are reluctant to pay extra cost-minimizing efforts to improve their efficiency outcomes. It indicates that the present of network constraint prevents the nodal pricing structure from providing useful price signals.

The findings of this chapter have important policy implications. In recent years, many countries and regions start to subsidize and install renewable energy generation sources. However, the progress of upgrading the transmission network is much slower compared to the speed of the installation of renewable sources. This chapter shows that, although renewable sources are beneficial to society in the long term, given the network constraint the new installations causes transmission congestion issues and leads fossil-fueled power plants in the surrounding areas to shirk in improving generation efficiency in the short term. Even if the regulator adopts more efficient pricing regulation, the network constraint could present as the obstacle for the efficiency gains. Future transmission network planning has to take such factors into consideration.

Appendices

Appendix A

Appendix for Chapter 2

A.1 Additional Data Cleaning Details

The cap-and-trade program in Texas was named as The Emissions Banking and Trading of Allowances Program. As the name suggests, it also allows for banking emission allowances and using the allowances left from the previous periods. However, the sample period only includes the first two years of this program and only two units use previous allowances in the control period 2004. Because of the rare banking behavior, this paper does not model the dynamic choice of banking allowances for future use.

A generating unit is a combination of fuel boiler and electricity generator, and such a unit together with the associated cooling system forms the basic unit of electricity production. Typically one boiler is uniquely linked to one generator. There are four generators in the sample linked to two identical boilers at the same time. All these four generators and associated boilers are not part of the cap-and-trade program. During the process of data cleaning, each group of the two identical boilers are aggregated into one boiler. The new boiler's capacity, annual emission and production is the summation of the original boilers.

There were also several types of generating units dropped from the sample due to data unavailability. A special type of generating units called peaking units are dropped. These units are usually set aside and only used during some unusual time in the summer to satisfy the unexpected high demand of electricity. Because these units are not included in the EIA or EPA database due to their low capacity, it is impossible to obtain their production data or other characteristics. Since their total emission is also very low and their allocated allowance is usually zero ton per year, dropping these units will not cause any substantial influence on the estimation results of this paper. Additionally, five coal-fired power plants in the sample have one additional auxiliary boiler. These boilers are not included in the EIA or EPA database, and it's impossible to identify which generator these boilers are linking with and their annual fuel usage. These boilers are also dropped from the sample. For Southwest NERC region, because the daily electricity prices were intermittent in the year 2001 due to the energy crisis, there was no reliable source of annual average price data in that year. Accordingly, the observations of the units located in Southwest NERC region in the control period 2001 were dropped.

A.2 Computational Procedure

The estimation of outside benefits and counterfactual simulations involve several assumptions and these assumptions are also used in the estimation of structural parameters. Firstly, it is assumed that pre-1971 units or post-1971 units within the same firm will participate into the cap-and-trade together at the same time. In the data the volunteer decisions for the post-1971 units owned by the same firm were the same. For the pre-1971 units, another assumption is that the pre-1971 units that were observed to be retired during 2001-2006 would choose whether to exit. The rest of the pre-1971 units are assumed to play the pure strategy of staying in the cap-and-trade markets under all scenarios. This assumption not only ensures that the number of the players in the participation game will not be too high to cause the curse of dimension, but also ensures that the first assumption is valid for the pre-1971 units. Given this assumption, the players in the participation game will be the groups of post-1971 units in each firm and the groups of pre-1971 units in each firm. Without these assumptions, there would be 2^{55} and 2^{150} potential combinations of participants in the west and east cap-and-trade markets respectively, which makes it impossible for computation. The data of 2004 control period are used for the estimation of the outside benefits and the counterfactual simulations.

The simulation steps are the following:

1. In each counterfactual policy regime listed in Table 2.13, there will be multiple players of the cap-and-trade participation depending on the number of firms in each emission market. Let the number of players be N_k in regime k, so that there would be 2^{N_k} combinations of participants in the cap-and-trade market and also 2^{N_k} equilibrium emission prices to be simulated. For each combination, the simulation begins by setting the allowance price to be \$1500/ton.

2. For each combination, given the set of the participants and the allowance price in the cap-and-trade market, the choices of abatement technology and production are simulated for each individual unit using the decision rules of the model. Each unit would choose the abatement technology with the highest second stage profits. Given the optimal technology choice, we can calculate the corresponding optimal production and emission for each unit. Then the sum of the total emissions of all cap-and-trade participating units is calculated.

3. If the sum of emission is higher than the total allowance allocation to the cap-and-trade participants (excluding exit units), the allowance price is increased by \$1/ton, otherwise lower by \$1/ton until the sum of the emission equal to the sum of allowance allocation for all participants. The sum of allowance allocation is calculated differently for the counterfactual simulation and the estimation in Section 6.4: (a) In the counterfactual simulations, it is assumed all cap-and-trade allowances to the participants excluding the exit units are used. (b) In the observed data, there are about 30k tons of allowances unused by the cap-and-trade participants. During the estimation of the outside benefits in Section 6.4, it is assumed that adjusting for the allocations to exit units, remaining units will never the extra allowances in the cap-and-trade. I assume there is extra transaction cost which increases the allowance price to prevent the cap-and-trade units to use all the allowances. The amount of unused allowances in each participation scenario equals to 30k tons minus the

sum of allowances to exit units.

4. After simulating the allowance prices, the equilibrium production and abatement choices of each unit in the cap-and-trade will be simulated. For units not participating in the cap-and-trade, if the unit is a post-1971 unit, the choices of this unit will be the same as in the data when they chose not to volunteer for the cap-and-trade program, or predicted by the model otherwise. Then, the profits of all the units in the second stage for each combination are calculated.

5. To calculate the profits of the units in the first stage, the same 50 draws of v in the section 6 are taken and the associated outside benefits are constructed for each unit using the estimated parameters for each draw. After that each unit's first stage payoffs are constructed for each draw.

6. For each draw of the S values, we can calculate the payoffs for each player in the participation game, which will be the sum of all the payoffs of the units owned by this player. Each player will choose the equilibrium probability of cap-and-trade participation.

7. For each draw of the S values, the expected welfare measures given these equilibrium probabilities are calculated. Finally, the averages of the welfare measures for these 50 draws are computed.

8. To compute the standard errors of the outside benefits parameters and the counterfactual results, I repeat the previous procedure 1-7 by ran-

domly draw the emission function, marginal cost and fixed costs parameters from the asymptotic distributions shown in Table 2.7 column (1) and Table 2.9 for 25 times.

There are four counterfactual regimes to be simulated in Table 2.13, and in these regimes the number of players in the participation game ranges from 3 to 13. In complete information games, there is concern about the multiplicity of equilibrium solutions and the associated equilibrium selection mechanism. During the simulation procedures, there always exists a strictly dominant strategy for each player so that the counterfactual games always have a unique equilibrium.

A.3 Robustness Checks

Table A.1 gives the test result on the model assumption that the individual fuel efficiency is not affected by the choices of abatement technologies. The dependent variable is the log of the ratio of generation over heat input. The coefficients on the technology dummies are all insignificant at 10% level, showing the assumption is valid.

Table A.2 gives the estimation results of alternative emission function specifications using unit-level fixed-effect regression. In the first two columns, the combination choice of Selective Catalyst Reduction with one of any other two categories is classified as the choice of Selective Catalyst Reduction. The last two columns do not impose such restriction in the classification. Table A.2

Table A.1: Test Results on Constant Fuel Efficiency Assumption

	log(q)
Combustion Modification	0.033
	(0.021)
Lower NOx Burner	-0.026
	(0.018)
Selective Catalyst Reduction	0.028
	(0.018)
sample selection	1.917**
	(0.973)
Year Dummies	Y
Unit Characteristics	Y
Observations	737
R-squared	0.200

Note: The result is estimated by the two-step Heckit methods controlling for part of the units retired during the sample period 2001-2004. The first step is to run the probit regression on the retirement decision on unit characteristics such as generator capacity, initial service year, wholesale market region and so on. Then the retirement probability is estimated for each unit. The second step is to add a variable which is the inverse Mills ratio of the predict retirement probability into the regression, denoted by 'sample selection'. The coefficients and standard errors are estimated via bootstrap in the second step (*** p¡0.01, ** p¡0.05, * p¡0.1).

shows that the coefficients do not change much when changing the classification criteria. The second and fourth columns relax the assumption that emission is linear in heat input. The coefficients of the log(heat input) are pretty close to one in these columns and the other coefficients are almost the same as with the linear assumption. Therefore, the linearity assumption is appropriate.

Table A.2: Robustness of Emission Function Specifications

VARIABLES	(1) log(e)-log(h)	(2) log(e)	(3) log(e)-log(h)	(4) log(e)
log(h)		1.030***		1.027***
		(0.015)		(0.015)
CM	-0.527***	-0.522***	-0.600***	-0.593***
	(0.075)	(0.075)	(0.075)	(0.075)
NB	-0.409***	-0.413***	-0.344***	-0.349***
	(0.066)	(0.066)	(0.066)	(0.066)
SCR	-1.561***	-1.571***	-1.397***	-1.405***
	(0.083)	(0.083)	(0.078)	(0.078)
Observations	737	737	737	737
R-squared	0.542	0.937	0.546	0.938

Note: This table presents the results of unit fixed-effect regression of the emission function using data from 2001 to 2004 for all units with positive emissions. Observations for exit units are dropped. Standard errors in parentheses (*** p¡0.01, ** p¡0.05, * p¡0.1). Constants are not reported. The first two columns classify the combination of SCR with other technologies as SCR.

In Table 2.7 the panel data for estimation are unbalanced because some units were retired during the sample period. This might cause the problem of sample selection. To address this problem, the column (2) in Table A.3 adds the initial service year as an additional variable to quantify the variable costs. Since only the old generations of units were likely to be retired, if the retirement selection is related to the variable costs, adding the initial service year into the regression equation will lead to significant changes in the estimation results of all coefficients. As Table A.3 shows, although the coefficient on the initial service year is negatively significant, the magnitude is sufficiently small. Moreover, the coefficients on the technology dummies and the intensity tax rate in column (2) are very similar to column (1). The only significant change in the estimation coefficients is the marginal fuel cost c_1. Therefore, there is no need to worry about sample selection in the estimation of variable costs.

Additionally, there is also concern that generating units using different type of fuel sources have different fuel costs, but in the paper I assume that the fuel costs for all units are the same. To check this assumption, in the column (3) of Table A.3 I add the dummy variable of natural gas indicator to examine whether fuel costs are different among generating units with different fuel types. The dummy variable natural gas is statistically insignificant, which supports the assumption in the paper.

In Table 2.14 of Section 2.5, the results are calculated based on the assumption that the electricity price is constant and all allowances are used by the participants in the cap-and-trade emission market. As stated in the simulation procedure above, in the observed data, there is a certain amount of allowances unused by the cap-and-trade participants. During the estimation of outside benefits in Section 6.4, it is assumed that adjusting for the allocations to exit units, remaining units will never use the extra allowances in the cap-and-trade. This assumption can be interpreted as there is extra

Table A.3: Variable Costs GMM Estimates ($/MMBtu)

	(1)	(2)	(3)
c_1	3.200***	5.178***	2.956***
	(0.097)	(0.003)	(0.513)
$c_2 * 10^6$	0.005***	0.006***	0.008
	(0.002)	(0.002)	(0.007)
Combustion Modification	0.226	0.221	0.243
	(0.220)	(0.220)	(0.240)
Lower NOx Burner	0.139	0.140	0.162
	(0.253)	(0.254)	(0.271)
Selective Catalyst Reduction	2.490***	2.492***	2.345***
	(0.513)	(0.513)	(0.497)
Intensity Tax Rate	3.586***	3.690***	6.469***
($/MWh*violation%)	(0.011)	(0.011)	(0.286)
Initial Service Year		-0.0001***	
		(0.000)	
Natural Gas Indicator			0.220
			(0.480)

Note: Column(1) in this table presents the estimation results of the equation (14) using 2001-2004 data for all units with positive emission and heat input. The number observations is 737. The number observations is 737. Observations for exit units are dropped. Column(2) presents the estimation results adding the initial service year. Column(3) presents the estimation results adding the dummy variable representing the generators using natural gas as primary fuel source. The reported scale of the coefficient c_2 is multiplied by 10^6. All the results are measured in $/MMBtu except for the intensity tax rate. Standard errors in parentheses (*** p¡0.01, ** p¡0.05, * p¡0.1).

transaction cost increasing the allowance price that prevents the units use all the allowances. During the counterfactual simulations this assumption is relaxed to assume the units will use all the allowances in the cap-and-trade. To examine the effect of this assumption on the counterfactual results, I compute the counterfactuals under different assumptions of allowance usage using the point estimates of cost parameters. Figure A.1 present the comparison results. Restriction on the usage of allowances only results in around 13k-17k tons reduction in aggregate emissions in each counterfactual regime. Because the revenues from emission allowance transactions in cap-and-trade account for only a small portion of total profits, the profits and exit rates remain almost unchanged. Note that the conclusion of this paper is drawn based on the difference between the welfare measures in each counterfactual regime. Therefore, the assumption on unused allowances does not affect the conclusions of this paper.

Another check is related to the assumption on electricity price. In the structural model, the electricity price is assumed to be constant. Because fossil-fueled generating units are usually price-setting units in the electricity market, this assumption might be too strong to hold. In this section, this assumption is relaxed and new results are estimated as a robustness check. Previous studies on the price elasticity of demand for electricity (Espey and Espey 2004; Neenan and Eom 2008) identified the long-run demand elasticity of electricity ranges from -0.3 to -3.26. This paper considers two scenarios, one with an elastic demand curve (elasticity=-2.5) and another an inelastic demand curve (elasticity=-0.8) to simulate for electricity prices by adjusting for the changes in total production. Changing the assumption on electricity price does not affect the estimation of the parameters in the second stage of the model, because these parameters are identified based on the observed production and electricity price. What will be affected are the estimation of the parameters in the first stage of the model and the counterfactuals. With the new assumed demand elasticity, in the simulation steps 2-3 of Appendix A.4, during each iteration of the simulation of emission price, the total production is re-calculated and the electricity price is adjusted by the changes in production relative to the observed total production of each wholesale market. The iterations stop when the total emissions equal to the total allowances allocation.

The results are also presented in Figure A.1. Changing the assumption on electricity price causes the aggregate emissions to be lower under the voluntary cap-and-trade regimes (1) and (2). The total emissions under the two-market mandatory cap-and-trade regulation are the highest, but the total emissions under the one-market mandatory cap-and-trade regulation are very similar to the voluntary regimes. Still, the changes in the total emissions are not significant. The aggregate profits and welfare are higher under the mandatory cap-and-trade regulations, which implies that the mandatory cap-and-trade regulation outperforms the mixed policy framework aggregately.

Figure A.1: Robustness Checks on Counterfactual Simulations

Note: This figure shows the comparison of counterfactual simulation results when relaxing the model assumptions with the results estimated under original assumptions. The results are measured on an annual basis by using 2004 data and the point estimates of the parameters of the emission function, marginal and fixed costs from previous subsections.

Bibliography

[1] Jose Miguel Abito. Agency costs in environmental regulation: Evidence from regulated electric utilities. Technical report, Working Paper, 2014.

[2] James B Bushnell. Adverse selection and emissions offsets. 2011.

[3] James B Bushnell, Stephen P Holland, Jonathan E Hughes, and Christopher R Knittel. Strategic policy choice in state-level regulation: The epa's clean power plan. *American Economic Journal: Economic Policy.*

[4] James B Bushnell and Catherine D Wolfram. Enforcement of vintage differentiated regulations: The case of new source review. *Journal of Environmental Economics and Management*, 64(2):137–152, 2012.

[5] H Ron Chan, Harrison Fell, Ian Lange, and Shanjun Li. Efficiency and environmental impacts of electricity restructuring on coal-fired power plants. *Journal of Environmental Economics and Management*, 81:1–18, 2017.

[6] In-Koo Cho and Hyunsook Kim. Market power and network constraint in a deregulated electricity market. *The Energy Journal*, pages 1–34, 2007.

[7] Steve Cicala. When does regulation distort costs? lessons from fuel procurement in us electricity generation. *The American Economic Review*, 105(1):411–444, 2014.

[8] Joseph A Cullen. Dynamic response to environmental regulation in the electricity industry. 2014.

[9] Evangelina Dardati. Pollution permit systems and firm dynamics: How does the allocation scheme matter? *International Economic Review*, 57(1):305–328, 2016.

[10] Lucas W Davis and Catherine Wolfram. Deregulation, consolidation, and efficiency: Evidence from us nuclear power. *American Economic Journal: Applied Economics*, 4(4):194–225, 2012.

[11] James A Espey and Molly Espey. Turning on the lights: a meta-analysis of residential electricity demand elasticities. *Journal of Agricultural and Applied Economics*, 36(01):65–81, 2004.

[12] Kira R Fabrizio, Nancy L Rose, and Catherine D Wolfram. Do markets reduce costs? assessing the impact of regulatory restructuring on us electric generation efficiency. *The American Economic Review*, 97(4):1250–1277, 2007.

[13] Carolyn Fischer. Combining rate-based and cap-and-trade emissions policies. *Climate Policy*, 3(sup2):S89–S103, 2003.

[14] Meredith Fowlie. Emissions trading, electricity restructuring, and investment in pollution abatement. *The American Economic Review*, 100(3):837–869, 2010.

[15] Meredith Fowlie and Nicholas Muller. Market-based emissions regulation when damages vary across sources: What are the gains from differentiation? Technical report, National Bureau of Economic Research, 2013.

[16] Meredith Fowlie, Mar Reguant, and Stephen P Ryan. Market-based emissions regulation and industry dynamics. *Journal of Political Economy*, 124(1):249–302, 2016.

[17] Lawrence H Goulder and Ian WH Parry. Instrument choice in environmental policy. *Review of environmental economics and policy*, 2(2):152–174, 2008.

[18] Catherine Hausman and Lucija Muehlenbachs. Price regulation and environmental externalities. 2016.

[19] Ali Hortaçsu, Seyed Ali Madanizadeh, and Steven L Puller. Power to choose? an analysis of consumer inertia in the residential electricity market. Technical report, National Bureau of Economic Research, 2015.

[20] Ali Hortaçsu and Steven L Puller. Understanding strategic bidding in multi-unit auctions: a case study of the texas electricity spot market. *The RAND Journal of Economics*, 39(1):86–114, 2008.

[21] Akshaya Jha. Regulatory induced risk aversion: Coal procurement at us power plants. Technical report, mimeo, 2015.

[22] Paul L Joskow and Jean Tirole. Transmission rights and market power on electric power networks. *The Rand Journal of Economics*, pages 450–487, 2000.

[23] Hyunsook Kim. *Transmission Capacity Effects on Market Inefficiencies in the California Electrical Network*. PhD thesis, University of Illinois at Urbana-Champaign, 2003.

[24] Jean-Jacques Laffont and Jean Tirole. *A theory of incentives in procurement and regulation*. MIT press, 1993.

[25] Paul Lehmann. Justifying a policy mix for pollution control: a review of economic literature. *Journal of Economic Surveys*, 26(1):71–97, 2012.

[26] Tanachai Limpaitoon, Yihsu Chen, and Shmuel S Oren. The impact of carbon cap and trade regulation on congested electricity market equilibrium. *Journal of Regulatory Economics*, 40(3):237–260, 2011.

[27] Adam Millard-Ball. The trouble with voluntary emissions trading: Uncertainty and adverse selection in sectoral crediting programs. *Journal of Environmental Economics and Management*, 65(1):40–55, 2013.

[28] Juan-Pablo Montero. Voluntary compliance with market-based environmental policy: Evidence from the us acid rain program. *Journal of Political Economy*, 107(5):998–1033, 1999.

[29] Juan-Pablo Montero. Optimal design of a phase-in emissions trading program. *Journal of Public Economics*, 75(2):273–291, 2000.

[30] Nicholas Z Muller. Linking policy to statistical uncertainty in air pollution damages. *The BE Journal of Economic Analysis & Policy*, 11(1), 2011.

[31] Bernard Neenan and Jiyong Eom. Price elasticity of demand for electricity: a primer and synthesis. *Electric Power Research Institute*, 2008.

[32] Randy A Nelson, Tom Tietenberg, and Michael R Donihue. Differential environmental regulation: effects on electric utility capital turnover and emissions. *The Review of Economics and Statistics*, pages 368–373, 1993.

[33] Karsten Neuhoff, Julian Barquin, Maroeska G Boots, Andreas Ehrenmann, Benjamin F Hobbs, Fieke AM Rijkers, and Miguel Vazquez. Network-constrained cournot models of liberalized electricity markets: the devil is in the details. *Energy Economics*, 27(3):495–525, 2005.

[34] Nitrogen Oxides. Why and how they are controlled. *Clean Air Technology Center (MD-12), US EPA Technical Bulletin No. EPA-456/F-99-006R. http://www. epa. gov/ttncatc1/dir1/fnoxdoc. pdf*, 1999.

[35] Thomas Roderick. Optimal transmission regulation in restructured electricity markets. 2013.

[36] Timothy J Stanton. Capacity utilization and new source bias: evidence from the us electric power industry. *Energy Economics*, 15(1):57–60, 1993.

[37] Arthur van Benthem, Suzi Kerr, et al. Optimizing voluntary deforestation policy in the face of adverse selection and costly transfers. Technical report, Motu Economic and Public Policy Research, 2010.

[38] Frank A Wolak. Regulating competition in wholesale electricity supply. In *Economic Regulation and Its Reform: What Have We Learned?*, pages 195–289. University of Chicago Press, 2014.

[39] Jay Zarnikau, Chi-Keung Woo, and R Baldick. Did the introduction of a nodal market structure impact wholesale electricity prices in the texas (ercot) market? *Journal of Regulatory Economics*, 45(2):194–208, 2014.

[40] Yiyuan Zhang. Market organization and productive efficiency: Evidence from the texas electricity market. Technical report, Agricultural and Applied Economics Association, 2016.