Transfer of Production and Emission Leakages: An Empirical Analysis on Phase II of the Acid Rain Program

A Thesis Submitted by

Hongxi Zhao

In partial fulfillment of the requirements of the degree of

Master of Science

in

Economics

Tufts University

May 2016

Advisor: Professor Gilbert E. Metcalf

Abstract

The Acid Rain Program was introduced by the U.S. Environmental Protection Agency in 1990 to control for sulfur dioxide emissions from the U.S. power sector. The program has been widely applauded for its environmental success. However, there has been little discussions on possible emission leakages issues in this program. The primary focus of my thesis is to test the potential for emission leakages problems in Phase II of the ARP: a transfer of production from regulated units to exempted units caused by compliance costs of the program. I conducted an empirical analysis on production performances of boilers in six U.S. northeastern states, using data from the EPA Air Market Program Database from 1999-2012. The empirical models used include difference-in-differences models testing behavioral changes of exempted and regulated units, and sulfur factor models testing exempted units' emission responses to the ARP. This paper supports the transfer of production hypothesis, but finds little evidences for the overall sulfur emission leakages. However, the findings in this paper indicate possible loopholes existing in cap-and-trade systems' regulations, and suggest that avoiding the substitutability between participants and non-participants is crucial for future environmental regulation designs to ensure the environmental effectiveness.

Acknowledgement

My deepest appreciation goes to my advisor, Professor Metcalf, for his instructions and support. I am extremely grateful for his patient tutoring, kind encouragements and rigorous teachings. Without him, this thesis would not be possible.

I also wish to express my gratitude to my reader, Professor Jack, for her advice and help.

Finally, I would like to thank my families for their love and support.

Table of Contents

I. Introduction

The Acid Rain Program (ARP), established under Title IV of 1990 Clean Air Act Amendments, is the first large-scale cap and trade program launched for pollution control. It is perceived as a regulation success from many perspectives: achieving environmental goals, decreasing costs of abatement, promotion of technologies and creating an efficient trading market (Burtraw et al 1998; Ellerman 2003; Frey 2013; Carlson et al 2000).

This program regulates SO_2 emissions at boiler/unit level. The U.S. Environmental Protection Agency (EPA) determines the applicability of the program based on individual unit's characteristics and monitors sulfur emissions from each unit. Power plants that own regulated unit(s) would hold allowance accounts at the EPA for the purpose of trading emission permits. At the end of each year, power plants need to prove that they own sufficient permits to offset total sulfur emissions from the regulated units they own.

Economists studying the effectiveness of this program mostly focus on electricity generating units under the ARP regulations. Regulated units are considered to have three principal compliance options: a). purchasing emission permits to compensate for their SO2 emissions; b). reducing sulfur dioxide emissions by introducing pollutants-removal technologies, such as scrubbers; c). switching to cleaner fuels with lower sulfur contents. Many past studies take them as the only three available choices for regulated units (Chan et al 2015; Ellerman 2000; Schmalensee et al 1998).

What has rarely been discussed is a loophole hidden in the program design; some generating units are exempted from complying with the ARP, allowing the electricity generation to be shifted from regulated units to these exempted units. There are two potential ways to achieve this. Power plants owning both regulated units and exempted units could attempt to utilize exempted units more, in order to save their compliance costs. Power plants owning only exempted units, who do not need to pay for their emission, would have relatively larger profit margin after the ARP launched compared with regulated units and be able to provide electricity at lower prices. Market forces can drive up demand for electricity from exempted units to substitute for electricity regulated units[1]. Due to data constraints, only the second mechanism is analyzed in this paper[2], as hypothesis for transfer of production.

This issue is extensively discussed as a part of emission leakage effects in analyzing the environmental effectiveness of an international carbon trading market[3]. Following the Kyoto Protocol, while some industrialized countries are prepared to cap their carbon emissions, there exists a concern that they may lose their comparative advantages in international markets. Production of goods would thus be carried out in non-complying countries, which is similar to this paper's transfer of production hypothesis. Emission leakages can be broadly defined as an increase in emissions from non-participants that offset emission reduction effort by participants of an

---

[1] This statement is true if regulated units and exempted units are substitutes of each other. Substitutability between electricity generating units highly depends on their capacity, operation costs and operation conditions. This will be discussed in detail in later sections.

[2] The database used in this paper contains only 12 power plants that own both exempted and regulated units.

[3] This paper only discusses what is often referred to as "short-term" emission leakages. Long-term emission leakages involve power plants migrating to non-regulated areas, which is commonly known as pollution haven hypothesis (Chen 2009). Since the ARP is a national program, the possibility of this migration is very small.

environmental program (Michelek,2015).

Ideas of emission leakages are brought into analyzing domestic regulations that are incomplete. Fowlie (2009) points out the underlying threat to program effectiveness from incomplete environmental regulations on the power sector, where electricity production can be transferred from regulated regions to non-regulated regions. Her definition of emission leakages is more specific: the difference between emissions under incomplete regulations and emissions under complete regulations. In her study, Fowlie also suggests that higher relative emission rates of non-participants are highly relevant for emission leakages to occur. If participants are cleaner than non-participants, the transfer of production may result in emission increases of non-participants exceeding emission reductions of participants.

The primary purpose of this paper is to test whether a transfer of electricity production from regulated units to exempted units occurred in six U.S. Northeastern states during Phase II of the Acid Rain Program. This paper also attempts to examine if exempted units in these six states have higher sulfur emission rates than regulated units. The rest of the paper is organized as follows: Section II covers the program's background and related literatures; Section III introduces empirical models for testing the transfer of production hypothesis. Section IV demonstrates results and tests the models.

Empirical results provide strong supportive evidences for the transfer of production hypothesis, indicating that exempted units became more utilized after the ARPII launched.

II. Background and Literature Review

II.A. Program Background

The Acid Rain Program (ARP) was introduced by George W. Bush's administration in 1989, as a part of Clean Air Act Amendments. It is designed to reduce SO2 emissions from the U.S. power sector using a cap and trade scheme. After revisions, the amendments passed both houses with high margins of votes in 1990. The 1990 Amendments granted the EPA the authority to implement an aggregate emission "cap" for all power generating units under ARP regulations. Allowances to emit can be purchased as sulfur emission permits, and traded between agencies and facilities. The EPA holds allowances auctions each year and awards highest bidders the amount of emission permits allowed under annual emission limits. From there, these permits go into a free trading market. Each regulated power generating facility must prove, at the end of each year, that it has sufficient permits to compensate for its annual SO2 emissions.

This cap-and-trade system have long been applauded as an efficient and cost-effective mechanism, by using economic incentives to achieve environmental goals. Phase I of the ARP launched in 1990, following nation-wide growing concerns on acid rain threats. The trading mechanism proved to be surprisingly efficient. The emission reduction goal was achieved significantly more swiftly than expected by the

EPA[4] (Schmalensee et al, 1998). Some facilities even participated in the program voluntarily, driven by economic incentives (Montero, 1999). By 2010, the program has successfully lowered its cap to 50% of the SO2 emission level in 1980.

There are two reasons for re-examining this extensively studied program. First, the sulfur emission allowance trading market re-opened in January of 2015 under Cross State Air Pollution Rule (CSAPR), following two years of legal halt on the program. However, unlike the ARP, the CSAPR is no longer a national program: only 28 states joined, and among them, only 23 states are required to reduce their SO2 emissions. Using Fowlie's definition, CSAPR is not a "complete" program, and may be vulnerable to emission leakages problems.

Second, cap-and-trade systems have been proposed as a regulating mechanism to control for greenhouse gases. This mechanism gains public acknowledgements through its promotion of compliance cost reductions. However, recent studies show that compared with carbon taxes, cap-and-trade programs do not bring significant fiscal or economic advantages (Pope and Owen, 2009). Furthermore, a revenue neutral carbon tax reform could have stronger economic and regulative performances than the trading scheme (Metcalf, 2009). Therefore, it is necessary to examine possible loopholes hidden in cap-and-trade policies for future regulation comparisons.

This paper focuses on Phase II of the ARP, which started on Jan 1, 2000. With this paper's concentrating on comparing exempted units and regulated units, detailed

[4] Researchers argued other reasons for the substantial emission decrease. The availability of coals with low sulfur contents during this period can be a significant contributor (Schmalensee and Stavins, 2012). Another theory is that Phase I participants had false expectations on future compliance costs, resulting in large amount of permits banking and emission reductions during Phase I (Ellerman, 2007).

regulations determining the applicability of the Acid Rain Program Phase II (ARPII) are listed below. A power generating facility usually contain multiple boilers (generating units) that are connected to electricity generators. Each boiler or unit is evaluated individually to determine if it meets exemption standards. Power plants, or power generating facilities, must demonstrate to the EPA reasons for their generating units to be exempted from the program at the beginning of each year, in order to avoid paying for their sulfur emissions.

Generally speaking, participation in the ARP is mandatory for all power plants with generating units of above 25MW nameplate capacity. Nameplate capacity is the maximum electricity production efficiency of a boiler/unit, determined by the unit's design when it is installed. In practice, there are multiple application standards that allow some units with nameplate capacity above 25MW to be exempted from the ARP. Below are detailed explanations.

Code of Federal Regulation Title 40 Part 72 (40 CFR 72) is the primary document EPA officials rely on to determine the applicability of the ARPII. For existing power plants (those operated before November 15, 1990), there are, in general, four cases where the power plants are exempted from Acid Rain Program regulations: (i). small generating unit, which could be a simple combustion turbine or a generator with 25MW or below nameplate capacity; (ii). power plants not generating electricity for sale;(iii). power plants which are bound by sale agreements to provide a considerable percentage of their electricity to a utility facility (around 15%), such as cogeneration facilities and some independent power production facilities; (iv). solid waste incinerators and other non-utility facilities. (40 CFR 72.6)

For new power plants, the exemptions include: (i). small generating units with total nameplate capacity of 25MW or below; (ii). "clean" facilities, those not burning coal or coal-derived fuels and those burning gaseous fuels with sulfur content lower than .05% (40 CFR 72.7).

In this paper, only generating units in utility facilities are analyzed. The database used contains no solid waste incinerators. Unfortunately, there is little information on what sale agreements each facility is bound to. Exempted units analyzed in this paper therefore fall into three exemption categories: 1). small units having nameplate capacity of 25MW or below; 2). units in generating facilities that are bound by sale agreements; 3). units in clean facilities.

II. B. Related Literatures

Studies on emission leakages began with debates over the Kyoto Protocol. Bohm (1993) first started to study reactions of non-participants to environmental programs without complete cooperation. His analysis was based on a hypothetical global demand-side carbon emission regulation. Bohm suggests that that participant countries' effort to reduce fuel consumption would be neutralized by increasing fuel consumptions from non-participant countries.

This idea was formalized by Felder and Rutherford (1993) and was widely known as carbon leakages: an emission increase from non-participants of an environmental regulations that at least partially offset emission reduction efforts by participants (Michelek, 2015). Felder and Rutherford's study focused on a hypothetical scenario where OECD countries unanimously curb their carbon emissions. Their paper indicates that there are two possible ways leading to the leakage: unregulated countries take over goods production that OECD countries forfeit, and hence emit more carbon than before; a decrease in fuel consumption from OECD countries brought down the world fuel prices and stimulate fuel consumption from unregulated countries. A later study by Burniaux and Martins (2012) showed that non-energy markets have relatively small impact on emission leakage rates.

However, in the power sector, the carbon leakage discussions became increasingly popular, and the analysis does not stop at discussions for demand-side regulations. There have been extensive discussions on carbon leakages of domestic cap-and-trade programs with incomplete cooperation: that is, only some regions are regulated. The discussions normally focus on the substitutability between regulated and non-regulated units.

In 2008, California government proposed an implementation of a cap-and-trade program to curb carbon emissions. However, more than 30% of California's electricity consumption was bought from out-of-state facilities (Chen, 2009), which means unregulated power generating units have the potential to replace regulated units within the state. Burtraw et al (2005) also expressed their concerns over possible carbon leakage issues through discussions on the sensitivity of carbon leakages to allowance allocation methods, since units holding more allowances can easily substitute other units' electricity production.

Fowlie (2009) later used a partial equilibrium model to simulate a possible carbon trading market in California. Her results suggest that emission leakages increase with the industry competitiveness and with pollution rates of unregulated units. Fowlie's findings are also consistent with the idea that substitutability makes higher emission leakages. Chen (2009) further improved the simulation models to study the mechanism and explored probable correlations between carbon leakages and the carbon emission allowances prices, which links carbon leakages to compliances costs of the power generating facilities.

This paper will attempt to test all of the three suggested contributors to emission leakages: substitutability, emission rates of exempted units and permit prices. Unfortunately, there is a severe missing data problem with the available database on sulfur emission values[5]. Given the fact that multiple previous literatures for different types of regulations have suggested that transfer of production from regulated to exempted units is a key step towards carbon leakages, this paper uses the existence of transfer of production to represent the potential for emission leakages.

A contribution of this paper is to study responses of a sub-group (exempted units of the ARP) that have rarely been discussed in the analysis of the Acid Rain Program's environmental effectiveness. Retirement and replacement of regulated units was brought up by Burtraw (1998) as a potential compliance choice, but was not discussed in detail. In general, economists measure the sulfur emission reductions from the acid rain program by using data on only the regulated units (Schmalensee et al 1998; Ellerman 2000, p323).

---

[5] In the available dataset, all data on pre-ARPII sulfur dioxide emissions for exempted units are 0's. This issue will be discussed again in later sections to explain its impacts on results analysis.

Exempted units from the ARP only constitutes less than 10% of the U.S. power generating sector. If 25MW was used as the cutoff value to solely determine the applicability of the program, based on generating units' information in 2000, maximum generating capacity (i.e. nameplate capacity) of all exempted units is only 4.50811E-05% of that of all U.S. generating units[6]. It is expected that possible emission increases from them would not have significant impact on the program's overall environmental effectiveness. However, it could shed lights on future policy designs.

III. Models and Data Descriptions

III.A. Transfer of Production Hypothesis and Difference-in-differences Models

If transfer of production occurs, it will be observed that exempted units become more utilized after the launch of the Acid Rain Program Phase II (ARPII). Due to the nature of the exemption standards as described in Section II.A, overall sizes of exempted units are smaller than those of regulated units. Therefore, exempted units are expected to generate less electricity in total, resulting in a gap of production between exempted and regulated units.

If the hypothesis on the transfer of production is true, this gap of production would narrow after the launch of the ARPII, as exempted units partially replace regulated ones for electricity generation. Therefore, the transfer of production

[6] Author's calculation based on EIA860a (2000) provided by U.S. Energy Information Administration.

hypothesis can be tested through testing changes on the gap of production following the launch of the ARPII. Under the assumption that there is no decrease in total electricity generation, a smaller gap after the ARPII launched means the total production from exempted units relatively increased compared with the production of from regulated units.

The test on changes of the gap of production is essentially a test on the difference between the ARPII's causal effects on electricity generation from exempted units and regulated units. This test requires comparisons of production both across time (before and after the launch of the ARPII), and across groups (a group of the exempted units and a group of the regulated units). A difference-in-differences (DD) model is chosen to execute the task. DD models can demonstrate program effects for each group at each time period using fixed effects, and estimate the difference between program effects using intersections of the fixed effects, which equates the change of the gap of production in this test. A DD model is therefore a good fit. The model is

$$\mathrm{Y_{ijt}} = \beta_0 + \delta_0\mathrm{ARPII_i} + \delta_1\mathrm{AFFECTED_{it}} + \delta_2\mathrm{ARPII_i} \times \mathrm{AFFECTED_{it}} + \beta_1\mathrm{CAPACITY_{ijt}} + \beta_2\mathrm{PFUEL} + \beta_3\mathrm{SFUEL} + \gamma_1\mathrm{YEAR_t} + \gamma_2\mathrm{STATE_j} + \gamma_3\mathrm{PROGRAMS_{ijt}} + \varepsilon_{ijt} \quad (1)$$

Three variables are used as dependent variables in this model: annual gross load, operation time and heat input, all of them being indicators of electricity generation. Gross load is each unit's annual total electricity generation, and serves as a measurement for overall output or production performances. Operation time and heat input are direct indicators for power plants' operation decisions, where operation time is the number of hours a unit operate each year and heat input is the total fuel input in an energy term. Unlike gross load, these two variables mainly reflect the willingness of power plants operators to utilize each unit.

The variables are in units i (boiler/generating units), j (states), and t (year). ARPII and AFFECTED are dummy variables for whether a unit is in the ARPII regulated group and whether the ARPII has affected the unit, respectively. ARPII is constant for each unit over time. This dummy variable would be 1 throughout the analyzed time period for any unit that participated in the ARPII, and be 0 otherwise. Unlike most standard DD models, AFFECTED varies not only over time but also across units, due to the fact that regulated units joined the ARPII in different years. For example, if a regulated unit joined the ARP in 2002, its AFFECTED would be 0 for years before 2002 and be 1 for years including and after 2002. Methods to determine these two variables will be explained in detail later in data descriptions.

The coefficient of AFFECTED, δ_1, measures the influence of the ARPII over all units' generation performances. Nevertheless, it cannot provide sufficient information on testing the transfer of production hypothesis; the hypothesis expects contrasting responses from regulated and exempted units, which would not be demonstrated by δ_1. A DD variable, the product of ARPII and AFFECTED, is needed to capture changes in generation performances both across units and over time. δ_2 (DD coefficient), the coefficient of the product, measures the difference-in-differences effect. The sign and magnitude of the DD coefficient indicate the trend and the size of changes on the generation gap between regulated and exempted units, and subsequently provide evidences for or against the transfer of production hypothesis.

CAPACITY is the annual operating capacity, calculated by $\frac{\text{Gross Load}_{ijt}}{Operation\ Time_{ijt}}$; this variable measures how much electricity a unit can generate in one hour, and represents a unit's electricity production efficiency. Production efficiency has not only high correlation with the total production, but also considerable impact over power plant operators' production choices: units with higher efficiency are expected to be favored by the operators for electricity generation, regardless of the ARPII's implementations. This variable is introduced into the model to control this effect.

PFUEL and SFUEL are dummy variables indicating if natural gas is the primary or secondary fuel of a boiler. In regions that this thesis focuses on, 74% of generating units burn natural gas as their primary fuel thanks to existing natural gas pipeline systems. Reasons for including them in this model will be explained later in data descriptions.

After the year 2008, in addition to the Acid Rain Program, the U.S. power sector faced more environmental programs under Clean Air Inter-State Rule and Regional Greenhouse Gas Initiative, regulating emissions of other air pollutants including nitrogen oxides and greenhouse gases. It is expected that these programs would also impact operation choices. Fixed effects for each program is therefore added to the model. Year fixed effect and state fixed effect are also accounted for in this model.

To further study the transfer of production hypothesis, another regression model is used to test the correlation between the transfer and the driving force of the transfer: compliance costs. Transfer of production is expected to occur since exempted units do not need to pay for their sulfur emissions. Therefore, the "cause" of any changes in production gap between exempted and regulated units is compliance costs of the ARPII.

To test this correlation, compliance costs are incorporated in the model. Computing compliance costs of a program can be very complicated[7]. This paper uses costs of abatement to represent compliance costs, which are the costs for power plants to reduce their sulfur emissions. Other miscellaneous costs that are relatively small, such as possible transaction costs of permits trading and related administrative costs, are omitted from this analysis. Theoretically speaking, in a cap-and-trade system like the Acid Rain Program, permit prices would equate marginal abatement costs (MAC)[8] (Tietenberg 1985, p20). A recent study suggests the efficiency in sulfur dioxide permit market allows permit prices to reflect MAC, especially after several years of operation (Hitaj and Stocking, 2016). When there is little barrier for permits trading, power plants are able to choose the most cost effective way to comply with the program, whether it is to purchase permits or to reduce emissions. Market forces would drive prices of these two options closer, since they are substitutes for each other, until permit price equate MAC. Therefore, SO2 emission permit price is used to represent the compliance costs. The model is:

$$\mathrm{Y_{ijt}} = \alpha_0 + \alpha_1 \mathrm{PRICE}_t + \sigma_0 \mathrm{ARPII_i} + \sigma_1 \mathrm{AFFECTED_{it}} + \sigma_2 \mathrm{ARPII_i} \times \mathrm{AFFECTED_{it}} + \sigma_3 \mathrm{PRICE}_t \times \mathrm{ARPII_i} \times \mathrm{AFFECTED_{it}} + \alpha_2 \mathrm{CAPACITY_{ijt}} + \alpha_3 \mathrm{PFUEL} + \alpha_4 \mathrm{SFUEL} + \vec{\theta} \bullet \mathbf{X} + \nu_{ijt} \qquad (2)$$

PRICE is the annual SO2 emission permit price. σ_3 measures how production gap

[7] For compliance costs computations, see Pope and Owen (2009), Carlson et al (2000) and Chestnut and Mills (2005).

[8] Marginal abatement cost (MAC) is the cost of reducing one additional unit of emission.

changed after the launch of the ARPII with changes in the permit price. If transfer of generation is, as described in the hypothesis, a result of the compliance costs, σ_3 (price DD coefficient) should be a statistically significant negative number. X in this model is a vector of fixed effect variables, as defined in Model (1).

III.B. Sulfur Emission Testing and Sulfur Factor Model

The DD models (1) and (2) can also be used to test the sulfur emission leakages problem. Past studies using DD models often adopt estimated counterfactual emissions as the baseline or control group data (Schmalensee et al 1998; Ellerman 2000, p323); Chan, Chupp et al, 2012). By using exempted units' data, emission responses of exempted units to the program would also be captured, providing a way to test if polluting behaviors of exempted units would change after the launch of the ARPII.

Similar to the analysis on transfer of production hypothesis, if sulfur emission leakages took place, there would have been a smaller sulfur emission gap between regulated and exempted units after the launch of the ARPII.

In practice, annual SO2 emissions is used as the dependent variable in the two regression models (1) and (2) to test the sulfur emission leakages hypothesis. δ_2 estimates possible changes in the emission gap and σ_3 the impact of compliance cost on such changes. If the hypothesis is true, both coefficients are expected to be significantly positive. For the purpose of this paper, (1) and (2) above will be used to test both hypotheses simultaneously.

As explained in previous sections, comparative emission rates of exempted units are expected to be a significant contributor to emission leakages. Emission increases from exempted units tend to have significant impacts if the exempted units are "dirtier" than regulated units. To measure the emission rate, sulfur factor is used, calculated by $\frac{SO_2\ Emissions_{ijt}}{Gross\ Load_{ijt}}$. Sulfur factor is the amount of sulfur dioxide emitted for a unit of gross load generated, which could control sulfur emissions variations brought by different electricity generation levels of exempted and regulated units. The model is:

$$SulfurFactor_{\mathrm{ijt}} = \omega_0 + \omega_1 \mathrm{PRICE}_t + \omega_2 \mathrm{ARPII_i} + \omega_3 \mathrm{PFUEL} + \omega_4 \mathrm{SFUEL} + \vec{\lambda} \bullet \mathbf{X} + \mathrm{u_{ijt}} \quad (3)$$

This model is used for analyzing post-ARPII data only, in order to compare exempted units to regulated units that have incentives to adopt abatement techniques. If exempted units are "dirtier" than regulated units, there should be a significant positive ω_2. This model not only tests the emission characteristics of the exempted units, but also demonstrates potential emission leakages at an individual level.

IV. Data Descriptions and Analysis

IV.A. Data Descriptions

The U.S. Environmental Protection Agency (EPA) and The U.S. Energy Information Administration (EIA) both provide detailed historic data at boiler (unit) level. The EPA only collects generation and emission data from units that have participated in at least one of EPA programs. The EIA datasets cover more generating units, but do not contain information on of units' participation status in the EPA programs. Furthermore, these two agencies identify boilers using distinctly different coding systems, which inhibits consolidation of their databases.

The main data source for this paper is the Air Market Program Database provided by the EPA, for years between 1999 and 2012. In order to obtain pre-ARPII (Acid Rain Program Phase II) data, I chose the database for the only program launched prior to the ARPII: The Ozone Transport Commission NOx Budget Program (OTC), which became effective in 1999. Although Phase II of the ARP was launched only one year later, not all units joined it in 2000. This database eventually provides 577 pre-ARPII observations and 1505 post-ARPII observations.

The OTC program is also a cap-and-trade program, designed to reduce nitrogen oxides emissions from the power sector. Participating states, instead of the EPA, were responsible for the regulation adoption, source identification and permits distributions. Eleven Northeastern states and several individual counties joined this program in 1999. In 2002, the OTC program was replaced by the NOx Budget Trading Program (NBP), which expanded the OTC program to include more states. Beginning in year 2009, the Clean Air Interstate Rule (CAIR) came effective, essentially taking over the role of the NBP. The CAIR, along with several other EPA programs, are accounted for as fixed effects in the DD models (1) and (2).

Since this paper focuses on Phase II of the ARP, units participating in the ARP Phase I are excluded from the database. All remaining units have participated in both the OTC and the NBP. This is to ensure that regulated units and exempted units being analyzed face similar policies and regulations. The final database covers six states: Connecticut, Massachusetts, New Jersey, New York, Pennsylvania, and Rhode Island.

Using the OTC and the NBP data calls for the need to control for different characteristics of the units burning natural gas. Combustion of natural gas has considerably higher nitrogen emission factor[9] than that of other fossil fuels. For electricity generation, natural gas has NOx factors of 140-550lbs/unit, depending on boilers' capacities. In comparison, coal has a NOx factor of 7.5-37lbs/unit; oil 24-67lb/unit; solid or liquid waste 3.8-20lb/unit. The OTC and NBP programs could make natural gas burning units less favorable, since their nitrogen emissions are substantially higher than other units. Therefore, dummy variables PFUEL and SFUEL are added to control for this effect.

If a unit joined the ARP any time after the year 2000, then it is considered to be an ARPII regulated unit. All other units in the database are treated as exempted units. Variable ARPII is 1 for all regulated units and 0 for all exempted units throughout the analyzed period.

It is less trivial to determine the variable AFFECTED. The time when the ARPII started to affect each unit must be determined first. The ARPII did not launch universally at the same time. Some units may have been granted exemptions at first but later joined the program due to changes in operation conditions. For a regulated unit, variable AFFECTED becomes 1 in the year the unit joined the ARP; for an exempted unit, the variable AFFFECTED depends on the county the unit locates in. After a regulated unit joined the ARP, all exempted units in the same county are considered to be affected by the program. Units in the same county normally share the same electricity dispatcher and end consumers, making them potential substitutes for

[9] Emission factor measures the environmental efficiency of power generation, by calculating the units of pollutants emitted for each MWh of power generated.

each other.

Table 1 shows the total number of regulated units that participated in the ARP each year in each county. Starting in the year that the number of units is no longer zero, every exempted unit in the same county is considered to be affected by the ARPII, with their corresponding variable AFFECTED becoming 1 thereafter.

The variable AFFECTED, which varies both over time and across units, provides a reasonable length of pre-treatment time period and a reasonable number of pre-treatment observations to conduct data analysis using DD models. Among over 600 units that was not affected by the ARP Phase I in 1999, only 8 units joined in 2000, followed by 11 units in 2001, 13 units in 2002, 8 units in 2003, 9 units in 2004, 6 units in 2005, 7 units in 2008, 1 unit in 2009 and 1 final unit in 2012. In total, there are 577 pre-treatment observations and 1505 post-treatment observations.

One final crucial variable to control for is the sulfur dioxide emission permit price (PRICE). The prices used in this paper are annual weighted average spot price for SO2 emissions permits awarded to winning bidders; the data is from the EIA's calculations based on the EPA's database[10].

IV.B. Comparability Issue

Before proceeding to the regression analysis, there are a few issues that need to be addressed. The first is whether exempted units are comparable to regulated units. Comparability determines the substitutability between exempted and regulated units, which is a critical determinant for the occurrence of transfer of production.

---

[10] Price levels are from "Average prices for spot sulfur dioxide emissions allowances at EPA auction set new lows", an article posed in *Today in Energy* by the EIA on May 11, 2011.

In this case, comparability can be measured by units' size. Since 25 MW nameplate capacity is used as a cutoff exemption standard, most exempted units have lower capacity than regulated units. The controlling variable CAPACITY can capture efficiency influences when the difference is small; however, the substitutability between units could decrease exponentially with greater capacity differences.

Considering a regulated unit with a generating capacity hundredfold of the capacity of an exempted unit, it can be considered irreplaceable by the exempted unit, because 100 hours of production from the exempted unit would be required to replace 1 hour of production from the regulated unit. In addition, there are structural differences between units with low and high capacities. Large units normally function as baseload electricity generators, providing electricity 24 hours at low costs; smaller units are generally peaking generators or shoulder generators, only generating electricity during peak periods or peak hours due to higher costs. Electricity grid operators have clear classifications for different types of generating units, in order to provide sufficient electricity cost-effectively. Generally speaking, units of different classifications are not substitutable for each other. Therefore, there may not be transfer of production if there is a large difference between capacities of exempted units and of regulated units.

Unfortunately, there are no data on Nameplate capacity (the maximum capacity that a unit can generate electricity at) in the EPA Air Market Program Database, and I have not identified an effective way to match the existing dataset to a database containing the nameplate capacity information. For the purpose of this paper, maximum operating capacity is used to compare units' production efficiencies. Operating capacity, calculated by $\frac{Gross\ Load}{Operation\ Time}$, is a measurement for a unit's generating capacity based on its real annual generation performances. The within standard deviation of operating capacity is 25.26MW. Compared with the sample mean value of 46.34MW, this is relatively small, indicating that the there is a small variation within each unit across time. Therefore, the maximum operating capacity is representative of the overall production efficiency level.

Maximum operating capacity is the highest operating capacity from 1999 to 2012. Although it may not equate a unit's highest possible capacity, it reflects the probable classification of a unit during the analyzed period, and such is a reasonable indicator for a unit's overall real-world production efficiency.

Table 2 shows summary statistics of key variables, including maximum operating capacity. Units with the highest maximum operating capacity are in the regulated group, suggesting that most baseload generators are probably regulated. A t-test on means of maximum operating capacity for both groups shows that regulated units have statistically significantly larger capacities than exempted units, as expected according to the ARP exemption standards. However, both groups' mean values of maximum operating capacities are in the capacity range for small units: on a national level, only less than 5% of units have nameplate capacity below 200MW[11]. The available data therefore yield inconclusive results for testing whether there exists a serious comparability issue. This issue will be addressed again in results and policy

[11] Calculated by author using EIA860a data.

implications section for further analysis. Additional analysis will be done through robustness tests in Section IV.B on whether the extreme values will be problematic for the regression models.

Another issue to be addressed here is the missing data problem of SO_2 values. Data for exempted unites on pre-ARPII SO_2 are missing for 400 exempted units, with all the recorded units having zero emissions. The missing data problem is expected to create biased coefficient estimation results of DD models (1) and (2) for sulfur dioxide emissions. The DD estimation results for SO_2 would be only suggestive in this case.

IV.C. Endogeneity Concerns

The DD models (1) and (2) may raise endogeneity concerns due to simultaneity of the ARPII applicability standards and dependent variables measuring production levels.

The first possible source of the simultaneity is nameplate capacity. As explained in Section II.A, a key factor to determine if a unit is subject to ARPII regulations is whether its nameplate capacity is above 25MW. Nameplate capacity, which is essentially a boiler's maximum generating capacity, also dictates a unit's generation efficiency. Therefore, it is expected to have high correlation with a boiler's operation performances and output level, which are measured by dependent variables used in the two models.

However, coefficient estimations should be consistent if there is no contemporary correlation between ARPII and nameplate capacity (which is in the error term in regression). Intuitively speaking, although the ARPII exemption standards are based on nameplate capacity, the capacity would not cause an endogeneity problem if an operator cannot actively change nameplate capacity to decide to participate in to drop from the program. In fact, this is what is expected, since nameplate capacity depends on boilers' designs when they first went on-line and is not easy to change. It is therefore necessary to test if nameplate capacity changed over time, especially during the periods that power plants operators are aware of the program and started to file for exemptions.

In order to examine possible changes in nameplate capacity, nationwide boilers' data is analyzed. Detailed regulations of Phase II of the Acid Rain Program was officially announced in the year 1993 and the program started to launch in year 2000. Table 4 shows a summary for boilers' nameplate capacity between these two years. There are 17% of units in total that made changes to their nameplate capacities. However, the fourth row and fifth row show that there are only in total 0.45% of units with nameplate capacity above 25MW lowered it to below 25MW. A further look at the data indicates that all of such units decreased their nameplate capacity to 0, which means they all went idle or retired afterwards. It is thus reasonable to assume that for most boilers, nameplate capacity is constant over time, and there is no endogeneity concern for the models.

Another possible source of the simultaneity comes from the exemption standard on purchasing agreements. As shown in the previous section, regulated units joined the ARP in different years. This could be the result of previously exempted units losing the exemption qualifications in later years. Among the three exemption standards applied to units in this database, as listed in program background, the only time-varying standard is for units being bound by purchasing agreements. Purchasing agreements essentially determine how much electricity generating units are providing to dispatchers and therefore have significant impacts over the output level, measured by dependent variables in Model (1) and (2).

No public information was found regarding reasons for exemptions or participations for the regulated units who joined the ARP after the year 2000. However, there does exist another possibility: voluntary participation motivated by economic incentives (Montero, 1999). If a unit was exempted due to its "clean" way of production, it could choose to join the program later voluntarily in hope to benefit from the permits trading market. In fact, among all 55 regulated units that joined after the year 2000, 48 burned natural gas as their primary fuel.

Unfortunately, there is no conclusive evidence for either possibility. It is necessary to note that the DD models may be subject to endogeneity problems in this case and provide biased estimation results.

IV.D. Common Trend Assumption

A key assumption of DD models is that pre-treatment control group and treated group share a common trend. It is crucial for obtaining unbiased estimates, because DD coefficients would capture differences caused by differing pre-treatment trends.

The common trend test used here is a chi-squared test on time trend coefficients before and after the launch of the ARPII, controlling all other covariates in the DD models. Time, t, is the number of years from or to the launch year of the ARPII for each unit, where negative numbers indicate the number of years before the program launched. Graph 1-4 illustrate the trends of Gross Load, Operation Time, Heat Input and SO_2 for exempted and regulated units, before and after the ARPII, with covariates controlled. In each graph, the dots indicate mean values of the variable at each time period with covariates controlled. The lines are best-fit lines to show the general trends. Since exempted units have all-zero pre-treatment SO_2 values, chi-squared test is not applicable on testing the SO_2 DD model.

From visual comparisons, in models with the production-measuring dependent variables, the exempted group and the regulated group seem to have very similar pre-treatment trends. Pre-treatment fitted values trends are roughly parallel to each other in Graph 1-3. Results from chi-squared tests in Table 4 also support the common trend assumption. All three variables pass chi-squared tests at 99% confidence level, with heat input and operation time also pass at 95% and 90% confidence level. Therefore, it is reasonable to assume common trend for these three depend variables.

Graph 4 indicates a disturbing trend where SO_2 emissions from exempted units increased after the ARPII launched. This is a potential threat to the program effectiveness as illustrated by the sulfur emission leakages hypothesis. What else is shown in this graph is that exempted and regulated units do not seem to share a common trend for sulfur dioxide emissions. This is probably one of the reasons why exempted units were rarely used as baselines in past DD analysis for the Acid Rain Program effectiveness. This would undermine the validity of coefficient estimations for SO_2 models, and will be addressed again in Section V.

V. Empirical Results

V.A. Regression results

For the purpose of displaying program effects in a clear and distinct way, variables in the regression models are all in log forms, so that coefficients reflect percentage changes. Level results are in Appendix for comparisons. Results of the two DD models are shown in Table 5 and Table 6.

There is a clear production shift observed from Table 5. The DD coefficient in column 3 suggests that the Acid Rain Program Phase II results in a strikingly 100% decrease in gross load gap between regulated and exempted units. DD coefficients in columns 1 and column 2 indicate that the change is caused by an active production choice of power plants operators, since there are similar decreases of 122% and 108% for gaps in heat input and operation time.

Narrower gaps in production performances between regulated and exempted units, as suggested by the negative DD coefficients, imply a decrease in the overall regulated units' electricity generation relative to the exempted units' generation. In the meantime, total power generation in the six states did not fall. Graph 5 shows total power production in each state from 1999-2012. Most states remained at the same level of power production while Pennsylvania experienced an increase. Therefore, given a relative decrease in production from regulated units, exempted units are expected to become more utilized after the launch of the ARPII, explicitly as described in the transfer of production hypothesis.

Coefficients on ARPII are significantly positive in Column 1-3 of Table 5, suggesting that regulated units overall are more utilized. The program may have shifted some production to exempted units, but regulated units are still the primary sources for electricity generation, having 106% more gross load than exempted units. Coefficients on operating capacity follow the expectation that units with higher production efficiencies are more preferable by operators.

Coefficients on natural gas dummy variables contradicts the intuition. The positive coefficients imply that units burning natural gases as primary or secondary fuels are actually more utilized. This could be the result of the availability of pipeline natural gases to Northeastern states, and/or compliance advantages from low sulfur and carbon content in natural gases.

The correlation between the transfer of production and its driving force is further examined in Table 6. PriceDD coefficient in Column 3 indicates that with 1% increase in the emission permit price, gross load would fall by 14%. Similar to Table 5, coefficient estimation results from heat input and operation time (Column 1 and 2) are also consistent with the gross load coefficient, suggesting that operators chose to shift to exempted units for power generation.

The focus now turns to the test on sulfur dioxide emissions. As discussed in the last section, due to violations of the common trend assumption, estimation results on SO2 DD coefficients may not be that reliable. Nevertheless, column 4 in Table 5 and Table 6 show DD coefficients consistent with past studies: regulated units emitted less sulfur dioxide relative to exempted units after the ARPII launched.

Model (3) compares emission rates of exempted units to that of regulated units.

Variables are also in their log forms in this case. Table 7 shows coefficient estimations from Model (3). Column 1 and 2 are estimation results from OLS estimation, while column 3 and 4 are results from GLS estimation. Column 2 and 4 excluded data from units that burn natural gas as their primary fuel source, in order to compare "dirty" units from both groups. Coefficient estimations on ARPII are consistent across different estimation methods, suggesting that regulated units actually have higher sulfur emission rate than exempted units. This could partially support the endogeneity analysis where most units were exempted for being in "clean" facilities.

For units burning other forms of fossil fuels (mainly oil and coal), the difference between the two groups seems larger. However, the level-form results in Appendix.C shows inconclusive results. Since the sample size for these particular type of units is small in this database, the coefficient estimations are not quite reliable.

V.B. Transfer of Production

Graph 6-8 further explores what happened to the regulated units after the ARPII launched. They show kernel density distributions of the three dependent variables: Heat Input, Operation Time and Gross Load of regulated units, before and after the launch of the program. Extreme values were removed in order to show clear distributive changes.

All three graphs illustrate density peaks shifting to lower levels of production or inputs after the ARPII was in effect, implying a general trend for regulated units to reduce their production activities. In addition, there are substantial increases in the number of idle units with zero gross load. This is probably linked to the production mechanism of power plants; power generators are difficult and costly to be turned on or off. Once operators decide to start utilizing a unit, it is inefficient to switch to other units. The observation of an increasing number of regulated units being put to idle is a clear signal that regulated units became less preferred after the ARPII.

The regression results and graphs consistently indicate that the transfer of generation occurred, as exempted units substituting regulated units after Phase II of the ARP launched. The scale of the shift is not only statistically significant but also quite large.

V.C. Robustness tests

From the model settings explained in Section III, there are two major concerns for the DD models' robustness: capacity comparability between the two groups and effects of other EPA programs that launched after year 2008.

Capacity comparability has a direct influence over the substitutability between regulated and exempted units. In Table 8, units with different ranges of operating capacities are selected for analysis. Only key coefficients, DD and price DD coefficients are reported. Coefficient estimations are quite close across different capacity ranges, implying that the models are not sensitive to capacity changes.

Table 9 conducts a similar robustness test on the EPA programs. Results in column 1-4 are from 1999-2007 data analysis where there were no other EPA programs involved other than ARP, OTC and NBP. Results in column 5-8 are from full database analysis with program fixed effects added to the models. There is a slight statistically significance change in coefficient estimation, but the sign and magnitude of the estimations do not vary significantly with dataset changes. Program fixed effects included in DD models well captured possible program influences and the coefficient estimations are robust in this case.

VI. Concluding Remarks

Results in this paper provide supportive evidences for the transfer of electricity production from regulated units to exempted units. Considering that there is a pending comparability issue between the exempted units and the regulated units analyzed in this paper, the transfer of production could be more significant if units from the two groups are more alike.

On a large scale, I could not find evidences on sulfur emission leakages themselves in the six states for this particular program. Due to the nature of the exemption standard for clean facilities, exempted units overall has lower sulfur factor than regulated units. Therefore, no evidence from this paper suggests that the environmental effectiveness of the ARP is being significantly compromised. This result is consistent with past studies. The lack of significance could be from the fact that this database contains only units from states where relatively cheap natural gas is available through pipelines. The Acid Rain Program's effects on exempted units' behaviors may be more obvious in more coal-dependent states, as suggested by the analysis on coal or diesel burning units.

From results in this paper, the existence of transfer of production itself posts a warning sign for future environmental regulation designers. To ensure the effectiveness of cap-and-trade programs, it is crucial to evaluate the substitutability of participants and non-participants, in avoidance of the transfer of production and possible emission leakages. For regional cap-and-trade programs, regulators need to consider the transfer across regional borders, as suggested by studies on RGGI (Fowlie 2009; Chen 2009). For national or international programs, exemption standards are the key to avoid the substitutions. The success of the Acid Rain Program in avoiding emission leakages comes from the fact that its exemption standards are set to: a). involve the absolute majority of polluting units and; b). ensure the categorical differences in generation abilities between exempted and regulated units.

Table 1. Number of Participating Units at County Level

# Units joined ARP	Year												
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Albany	0	0	0	0	0	0	0	0	1	1	1	1	1
Bristol	0	0	0	0	0	1	1	1	1	1	1	1	1
Camden	0	0	1	1	1	1	1	1	1	1	1	1	1
Cattaraugus	0	1	1	1	1	1	1	1	1	1	1	1	1
Clinton	0	0	0	0	0	0	0	0	0	2	2	2	2
Erie	0	2	3	3	3	3	3	3	3	3	3	3	3
Essex	0	2	2	2	3	3	3	3	3	2	2	2	2
Genesee	0	0	0	1	1	1	1	1	1	1	1	1	1
Gloucester	0	0	2	2	2	2	2	2	2	2	2	2	2
Hampden	0	0	0	0	0	2	2	2	2	2	2	2	2
Hartford	0	0	1	1	1	1	1	1	1	1	1	1	1
Jefferson	1	1	1	1	1	1	1	1	1	1	1	1	1
Middlesex	0	1	1	1	4	5	5	5	5	5	5	5	5
Nassau	0	0	0	0	2	2	2	2	2	2	2	2	2
Niagara	1	1	1	1	1	1	1	4	4	4	4	4	4
Norfolk	0	0	0	0	0	2	2	2	2	2	2	2	2
Northampton	0	0	0	0	0	0	0	0	0	0	0	0	1
Northumberland	0	0	1	1	1	1	1	1	1	1	1	1	1
Oneida	0	0	0	1	1	1	1	1	1	1	1	1	1
Onondaga	3	3	3	5	5	5	5	5	5	5	5	3	3
Oswego	0	0	1	2	6	6	6	6	6	5	5	5	5
Providence	0	0	1	1	1	1	1	1	5	5	5	5	5
Rensselaer	0	2	2	2	2	2	2	2	2	2	2	2	2
Saint Lawrence	0	0	0	3	3	3	3	3	3	3	3	3	3
Salem	1	1	1	1	1	1	1	1	1	1	1	1	1
Saratoga	0	0	1	1	1	1	1	1	1	1	1	1	1
Schuylkill	1	1	1	1	1	1	1	1	4	5	5	6	6
Worcester	0	0	0	0	0	0	0	0	0	1	1	1	1
Wyoming	0	0	1	1	1	1	1	1	1	1	1	1	1
Total	7	15	25	33	43	49	49	52	60	62	62	61	62

Data source: calculated by author

Note: The table shows the total number of regulated units that have joined the ARP each year in each county. The decrease in the number is caused by missing observations rather than units leaving the program.

Table 2. Descriptive Summary Statistics for All Units

	ARPII Exempted Units					ARPII Regulated Units				
	Mean	Standard Deviation	Minimum	Maximum	Number of Observations	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Heat Input (1000 GBtu)	852.7	2,050.5	0.2	18,104.9	1,527	2,939.3	3,138.2	0.5	15,659.2	864
Operation Time (hours)	1,896.2	3,052.6	0.0	8,779.3	1,610	3,616.2	3,047.2	0.0	8,766.9	911
Gross Load (1000 GWh)	52.6	158.8	0.0	1,579.2	1,251	320.6	866.8	0.0	13,936.8	831
SO_2 Emissions (tons)	35.1	145.3	0.0	1,327.2	512	68.3	229.8	0.0	1,938.7	561
Capacity (MW)	28.2	22.3	0.0	189.2	1,251	73.6	104.6	12.6	1,668.9	831
Permit Price (US dollars)	N/A	N/A	N/A	N/A	N/A	264.3	253.4	0.0	883.1	912
Sulfur Factor (tons/1000GWh)	0.21	0.64	0.00	3.95	433.00	0.16	0.54	0.00	3.21	543.00
Maximum Operation Capacity (MW)	40.9	34.3	0.0	189.2	1,409	105.3	205.6	24.0	1,668.9	870

Data source: EPA Air Market Program Database 1999-2012

Note: This table is to compare statistics of the two groups; it is expected that overall regulated units generate more than exempted units, as shown in this table; permit prices only apply to regulated units.

Table 3. A Summary of Nameplate Capacity

	1994	1995	1996	1997	1998	1999	2000
Total number of units	10060	10060	10060	10060	10060	10060	10060
Number of units with changed nameplate capacity	263	285	291	496	262	67	85
Percentage of untis with changed nameplate capacity	2.61%	2.83%	2.89%	4.93%	2.60%	0.67%	0.84%
Number of units changing to nameplate capacity <25MW	0	6	1	37	2	0	0
Percentage of units changing to nameplate capacity <25MW	0.00%	0.06%	0.01%	0.37%	0.02%	0.00%	0.00%

Data source: calculated by author based on eia-860A 1993-2000 from U.S. Energy Information Administration

Note:Data of 1993 is not shown here because this table provides information mainly on units changing their nameplate capacity.

Table 4. Chi-squared Test Results for Common Trend Assumption

	Exempted Group Time Coefficient	Regulated Group Time Coefficient	Chi-squared	P-value
Heat Input	-605,829***	-526,089***	0.16	0.687
	(77,532)	(105,535)		
Operation Time	-208.9	-514.9***	1.02	0.3135
	(148.2)	(69.16)		
Gross Load	-57,745***	-23,154***	4.16	0.0413
	(7,034)	(4,792)		

Standard errors in parentheses

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

Note: Time coefficients are obtained from OLS regressions of dependent variables on time (the number of years to or since the ARP) and other covariates in DID models; p-values are from chi-squared tests for the null hypothesis: exempted group time coefficient is equal to regulated group time coefficient.

Table 5. Regression Coefficient Estimation Results for DID Models

VARIABLES	(1) log(Heat input)	(2) log(Operating time)	(3) log(Gross load)	(4) log(SO_2 emissions)
ARPII	1.798***	1.509***	1.065***	0.929***
	(0.144)	(0.130)	(0.118)	(0.256)
Affected	0.00630	-0.000698	0.0174	0.467***
	(0.147)	(0.133)	(0.121)	(0.164)
DID	-1.220***	-1.082***	-1.002***	-0.482*
	(0.157)	(0.141)	(0.128)	(0.259)
log(Operating Capacity)	1.240***	0.729***	2.180***	0.300***
	(0.0525)	(0.0476)	(0.0433)	-0.0507
Natural gas as primary fuel	1.143***	1.218***	1.567***	-1.480***
	(0.0999)	(0.0903)	(0.0821)	(0.0943)
Natural gas as secondary fuel	0.964***	0.702***	0.911***	1.403***
	(0.185)	(0.168)	(0.153)	(0.166)
Constant	6.409***	2.598***	1.223***	-1.375***
	(0.238)	(0.215)	(0.195)	(0.442)
Observations	2,072	2,082	2,082	976
R-squared	0.597	0.524	0.783	0.591

Standard errors in parentheses

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

Note: The table does not report coefficients on fixed effects; the log values of varialbes var are calculated by log(1+var) in order to prevent undefined log values brought by level values of 0.

Table 6. Regression Coefficient Estimation Results for DID Models with Permit Price

VARIABLES	(1) log(Heat input)	(2) log(Operating time)	(3) log(Gross load)	(4) log(SO_2 emissions)
log(Price)	-0.203*	-0.215**	-0.245**	-0.199
	(0.116)	(0.105)	(0.0953)	(0.121)
ARPII	1.796***	1.508***	1.064***	1.008***
	(0.143)	(0.129)	(0.117)	(0.254)
Affected	0.0450	0.0386	0.0506	0.382**
	(0.147)	(0.133)	(0.121)	(0.163)
DID	-0.453*	-0.309	-0.350*	-1.112***
	(0.241)	(0.218)	(0.198)	(0.285)
log(Price)*DID	-0.165***	-0.166***	-0.140***	0.145***
	(0.0394)	(0.0358)	(0.0326)	(0.0288)
log(Operating Capacity)	1.238***	0.726***	2.178***	0.299***
	(0.0523)	(0.0474)	(0.0431)	(0.0500)
Natural gas as primary fuel	1.157***	1.231***	1.578***	-1.581***
	(0.0996)	(0.0899)	(0.0818)	(0.0952)
Natural gas as secondary fuel	1.009***	0.747***	0.949***	1.424***
	(0.185)	(0.167)	(0.152)	(0.164)
Constant	7.486***	3.742***	2.527***	-0.142
	(0.646)	(0.586)	(0.533)	(0.575)
Observations	2,072	2,082	2,082	976
R-squared	0.600	0.529	0.785	0.602

Standard errors in parentheses

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

Note: The table does not report coefficients on fixed effects; the log values of varialbes var are calculated by log(1+var) in order to prevent undefined log values brought by level values of 0.

Table 7. Regression Coefficient Estimation Results of Sulfur Factor Model in Log Forms

Log(Sulfur Factor)	(1) OLS	(2) OLS	(3) GLS	(4) GLS
log(Permit Price)	0.0352	-0.0662**	0.0352	-0.0662**
	(0.0261)	(0.0312)	(0.0258)	(0.0298)
ARPII	0.0928***	0.247**	0.0928***	0.247***
	(0.0196)	(0.0959)	(0.0194)	(0.0915)
Natural gas as primary fuel	-0.542***		-0.542***	
	(0.0251)		(0.0248)	
Natural gas as secondary fuel	0.149***		0.149***	
	(0.0455)		(0.0449)	
Constant	0.202***	0.555***	0.202***	0.555***
	(0.0345)	(0.0671)	(0.0341)	(0.0641)
Including units burning natural gas	Y	N	Y	N
Observations	916	203	916	203
R-squared	0.514	0.880		
Number of identification			136	34

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

Note: This table contains coefficient estimations for Model (3) in log forms; the first two columns are OLS estimation results and the last two columns are GLS estimation results; Column 2 and 4 shows results of analyzing on units not burning natural gas.

Table 8. Robustness Test on Operating Capacity

	(1)	(2)	(3)	(4)	(5)	(6)
log(Gross load)						
DID	-1.177***	-1.028***	-1.041***	-0.586***	-0.375*	-0.394**
priceDID				-0.125***	-0.140***	-0.139***
log(Heat input)						
DID	-1.303***	-1.283***	-1.263***	-0.620**	-0.548**	-0.502**
priceDID				-0.144***	-0.158***	-0.164***
log(Operating time)						
DID	-1.188***	-1.110***	-1.106***	-0.506**	-0.353	-0.336
priceDID				-0.144***	-0.163***	-0.166***
log(SO2 emissions)						
DID	-0.164	-0.473*	-0.482*	-0.897***	-1.123***	-1.112***
priceDID				0.158***	0.149***	0.145***
Capacity: (0-100MW)	Y	Y	Y	Y	Y	Y
Capacity: (100-200MW)	N	Y	Y	N	Y	Y
Capacity: (200-500MW)	N	N	Y	N	N	Y

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

Note: The table reports DID coefficient and PriceDID coefficient for each dependent variable given different capacity ranges; the first three columns are for Model (1) and the last three columns are for Model (2); the log values of varialbes var are calculated by log(1+var) in order to prevent undefined log values brought by level values of 0.

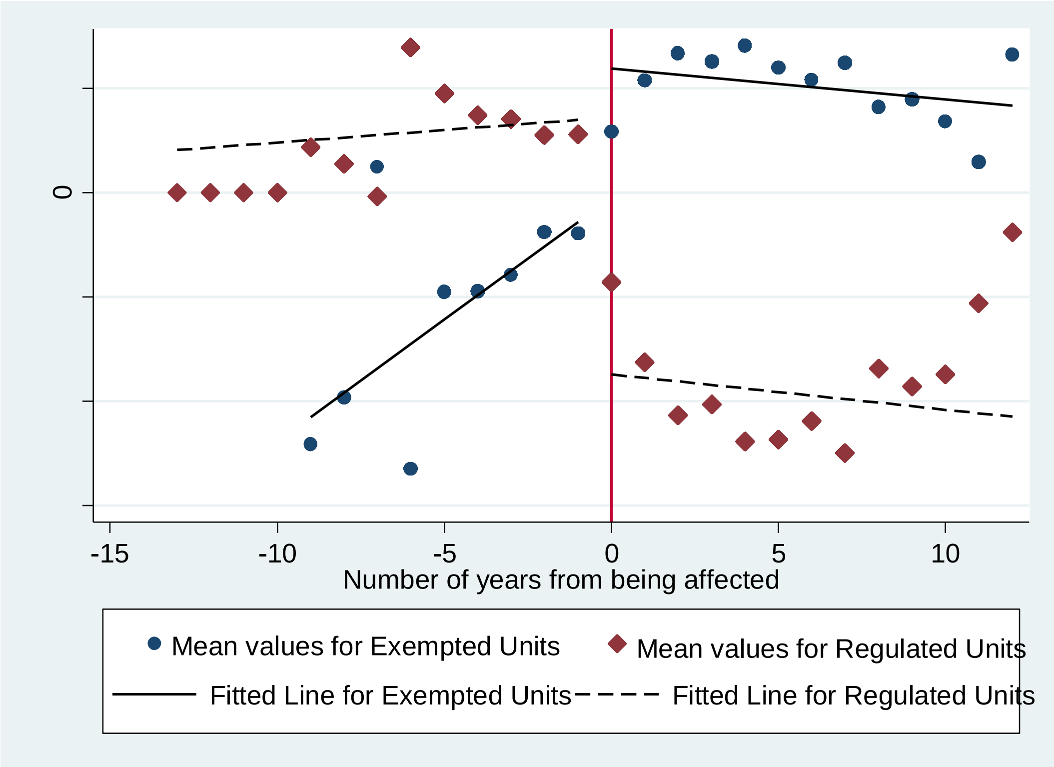
Table 9. Robustness Test on Programs Participation

VARIABLES	(1) log(Heat input)	(2) log(Operation time)	(3) log(Gross load)	(4) log(SO2 emissions)	(5) log(Heat input)	(6) log(Operation time)	(7) log(Gross load)	(8) log(SO2 emissions)
Model1								
DID	-1.159***	-1.068***	-0.814***	1.194***	-1.070***	-1.004***	-0.810***	-0.473*
Model2								
DID	0.425	0.848	2.026**	0.757	-0.453*	-0.309	-0.350*	-1.112***
priceDID	-0.282	-0.328**	-0.495***	0.0719	-0.165***	-0.166***	-0.140***	0.145***
Including 2008-2012 Data	N	N	N	N	Y	Y	Y	Y

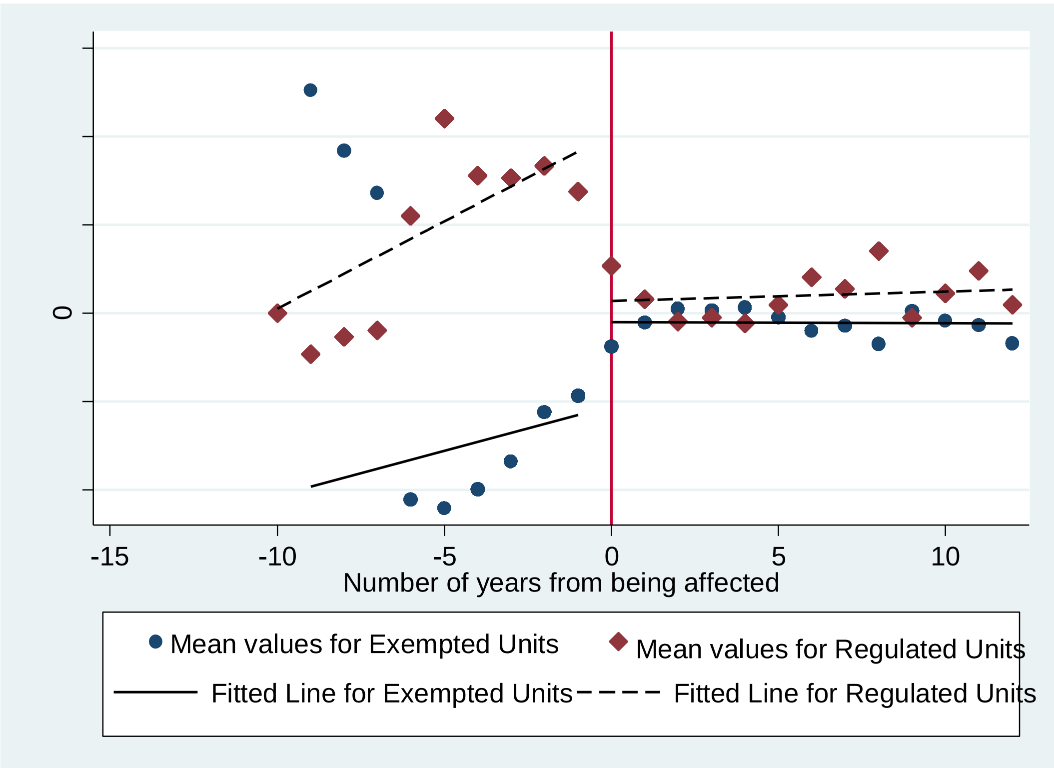
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The table reports DID coefficients and PriceDID coefficients with observations in different time ranges; the log values of varialbes var are calculated by log(1+var) in order to prevent undefined log values brought by level values of 0.

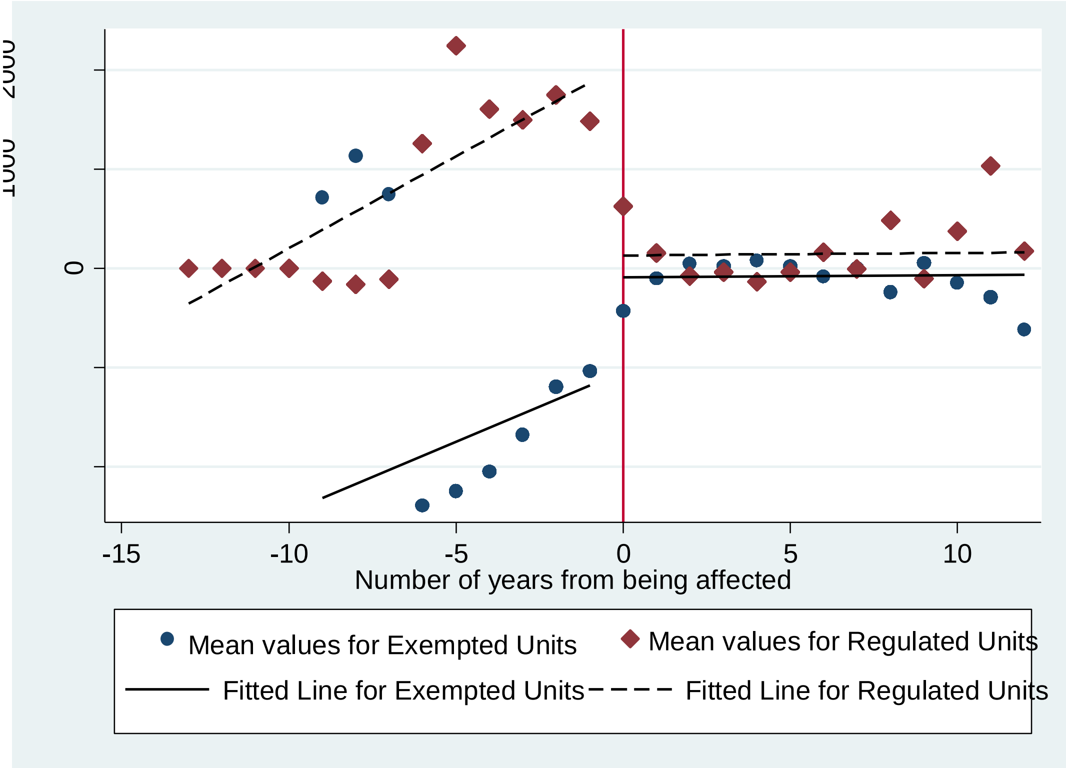
Graph 1. Trends of Gross Load with Covariates Controlled



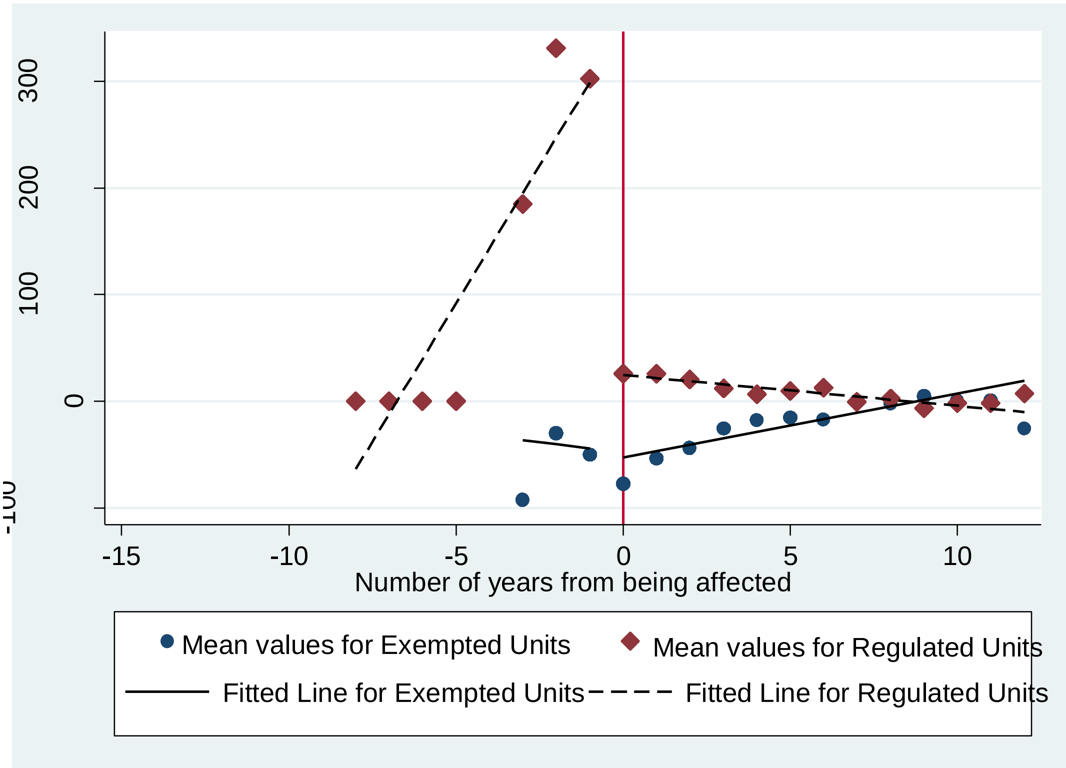
Graph 2. Trends of Heat Input with Covariates Controlled



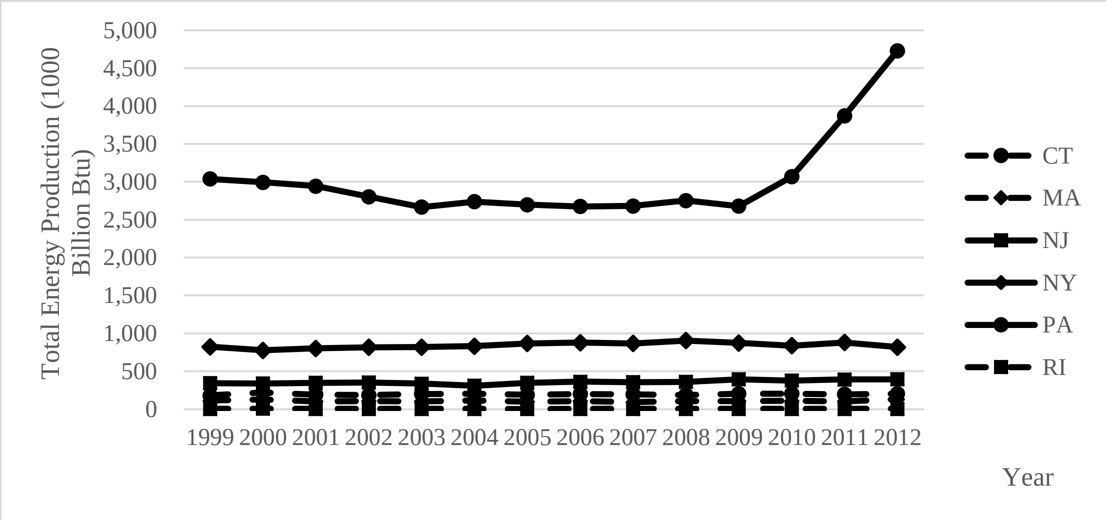
Graph 3. Trends of Operation Time with Covariates Controlled



Graph 4. Trends of SO_2 Emissions with Controlled Covariates

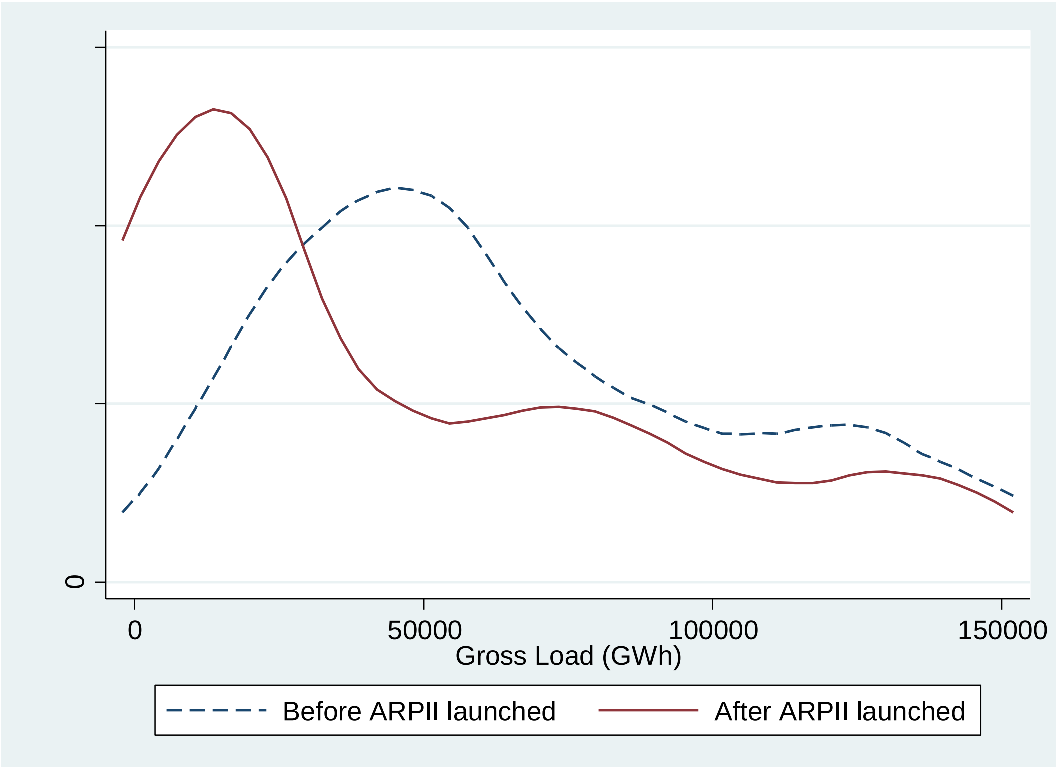


Graph 5. Total Energy Production across States

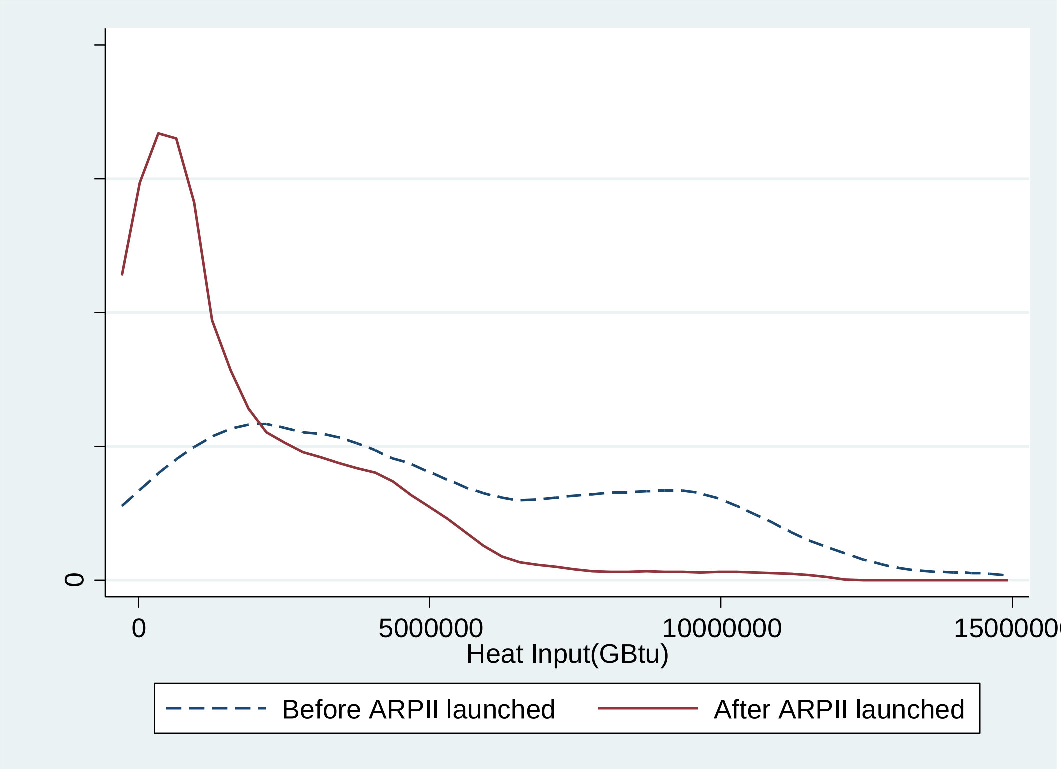


Data Source: The U.S. Energy Information Administration

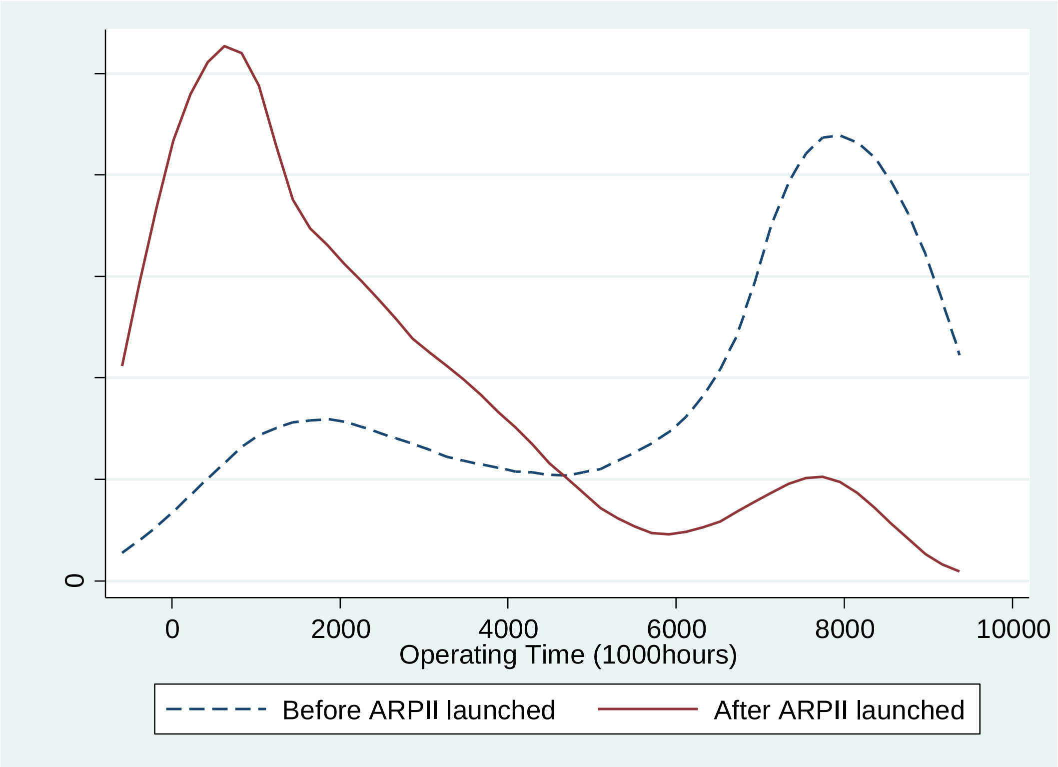
Graph 6. Kernal Density Distributions of Gross Load



Graph 7. Kernal Density Distributions of Heat Input



Graph 8. Kernal Density Distributions of Operation Time



Bibliography

Bohm, Peter. "Incomplete International Cooperation to Reduce CO2 Emissions: Alternative Policies." *Journal of Environmental Economics and Management* 24.3 (1993): 258-71.

Burniaux, Jean-Marc, and Joaquim Oliveira Martins. "Carbon Leakages: A General Equilibrium View." *Economic Theory* 49.2 (2012): 473-95.

Burtraw, Dallas, et al. "Costs and Benefits of Reducing Air Pollutants Related to Acid Rain." *Contemporary Economic Policy* 16.4 (1998): 379-400.

Burtraw, Dallas, et al. "Costs and Benefits of Reducing Air Pollutants Related to Acid Rain." *Contemporary Economic Policy* 16.4 (1998): 379-400.

Carlson, Curtis, et al. "Sulfur Dioxide Control by Electric Utilities: What are the Gains from Trade?" *Journal of Political Economy* 108.6 (2000): 1292-326.

Chan, Gabriel, et al. "The so^sub 2^ Allowance-Trading System and the Clean Air Act Amendments of 1990: Reflections on 20 Years of Policy Innovation." *National Tax Journal* 65.2 (2012): 419-52.

Chan, H. Ron, et al. *The Market for Sulfur Dioxide Allowances: What have we Learned from the Grand Policy Experiment?*. w21383 Vol. , 2015.

Chen, Yihsu. "Does a Regional Greenhouse Gas Policy make Sense? A Case Study of Carbon Leakage and Emissions Spillover." *Energy Economics* 31.5 (2009): 667-75.

Chestnut, Lauraine G., and David M. Mills. "A Fresh Look at the Benefits and Costs of the US Acid Rain Program." *Journal of environmental management* 77.3 (2005): 252-66.

Ellerman, A. Denny, and Juan-Pablo Montero. "The Efficiency and Robustness of Allowance Banking in the U.S. Acid Rain Program." *The Energy Journal* 28.4 (2007): 47-71.

Ellerman, A. Denny. *Lessons from Phase 2 Compliance with the U.S. Acid Rain Program.*, 2003.

Ellerman, A. Denny, and Juan-Pablo Montero. "The Declining Trend in Sulfur Dioxide Emissions: Implications for Allowance Prices." *Journal of Environmental Economics and Management* 36.1 (1998): 26-45.

Ellerman, A. Denny. *Markets for Clean Air :The U.S. Acid Rain Program*. Cambridge, U.K.; New York: Cambridge University Press, 2000.

Felder, Stefan, and Thomas F. Rutherford. "Unilateral CO2 Reductions and Carbon Leakage: The Consequences of International Trade in Oil and Basic Materials." *Journal of Environmental Economics and Management* 25.2 (1993): 162.

Fowlie, Meredith L. "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage." *American Economic Journal: Economic Policy* 1.2 (2009): 72-112.

Goulder, Lawrence H. "Markets for Pollution Allowances: What are the (New) Lessons?" *The Journal of Economic Perspectives* 27.1 (2013): 87-102.

Hitaj, Claudia, and Andrew Stocking. "Market Efficiency and the U.S. Market for Sulfur Dioxide Allowances." *Energy Economics* 55 (2016): 135-47.

Joskow, Paul L., Richard Schmalensee, and Elizabeth M. Bailey. "The Market for Sulfur Dioxide Emissions." *The American Economic Review* 88.4 (1998): 669-85.

Joskow, Paul L., and Richard Schmalensee. "The Political Economy of Market‐Based Environmental Policy: The U.S. Acid Rain Program." *The Journal of Law & Economics* 41.1 (1998): 37-84.

Metcalf, Gilbert E. "Designing a Carbon Tax to Reduce U.S. Greenhouse Gas Emissions." *Review of Environmental Economics and Policy* 3.1 (2009): 63-83.

Michalek, Gabriela, and Reimund Schwarze. "Carbon Leakage: Pollution, Trade Or Politics?" *Environment, Development and Sustainability* 17.6 (2015): 1471-92.

Montero, Juan‐Pablo. "Voluntary Compliance with Market‐Based Environmental Policy: Evidence from the U. S. Acid Rain Program." *Journal of Political Economy* 107.5 (1999): 998-1033.

Napolitano, Sam, et al. "The U.S. Acid Rain Program: Key Insights from the Design, Operation, and Assessment of a Cap-and-Trade Program." *The Electricity Journal* 20.7 (2007): 47-58.

Pope, Jeff, and Anthony D. Owen. "Emission Trading Schemes: Potential Revenue Effects, Compliance Costs and overall Tax Policy Issues." *Energy Policy* 37.11 (2009): 4595-603.

Resources for, the Future, Carolyn Fischer, and Stephen W. Salant. *Alternative Climate Policies and Intertemporal Emissions Leakage: Quantifying the Green Paradox.*, 2012.

Schmalansee, Richard, et al. "An Interim Evaluation of Sulfur Dioxide Emissions Trading." *The Journal of Economic Perspectives* 12.3 (1998): 53-68.

Schmalensee, Richard, and Robert N. Stavins. "The SO_2 Allowance Trading System: The Ironic History of a Grand Policy Experiment." *The Journal of Economic Perspectives* 27.1 (2013): 103-21.

Stavins, Robert N. "What can we Learn from the Grand Policy Experiment? Lessons from SO_2 Allowance Trading." *The Journal of Economic Perspectives* 12.3 (1998): 69-88.

Tietenberg, Thomas H. *Emissions Trading, an Exercise in Reforming Pollution Policy*. Washington, D.C.: Resources for the Future, 1985.

Appendix. Level-Form Regression Coefficient Estimation Results

Table A. Regression Coefficient Estimation Results for DID Models in Level Form

VARIABLES	(1) Heat Input (1000GBtu)	(2) Operating Time (1000hours)	(3) Gross Load (1000GWh)	(4) SO2 (tons)
ARPII	3,162***	3,203***	174.4***	134.8***
	(191.9)	(202.7)	(46.32)	(26.41)
Affected	-84.40	387.1*	134.5***	21.89
	(196.3)	(208.2)	(47.57)	(16.94)
DID	-3,110***	-3,414***	-414.2***	-88.68***
	(209.3)	(221.0)	(50.51)	(26.71)
log(Operating Capacity)	1,060***	1,067***	353.3***	7.496
	(70.13)	(74.46)	(17.02)	(5.220)
Natural gas as primary fuel	-603.2***	-286.4**	-143.6***	-111.9***
	(133.5)	(141.3)	(32.29)	(9.717)
Natural gas as secondary fuel	-478.9*	-123.2	-131.7**	215.1***
	(247.4)	(262.9)	(60.07)	(17.11)
Constant	-2,751***	-1,947***	-960.0***	-75.24*
	(317.5)	(336.4)	(76.88)	(45.51)
Observations	2,072	2,082	2,082	976
R-squared	0.418	0.438	0.296	0.509

Standard errors in parentheses

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

Note: The table reports regression results from Model (1) in level forms; the coefficient estimations here have same signs as in the log-form results.

Table B. Regression Coefficient Estimation Results for DID Models with Permit Price in Level Form

	(1)	(2)	(3)	(4)
		Operating		
	Heat Input	Time	Gross Load	SO2
VARIABLES	(1000GBtu)	(1000hours)	(1000GWh)	(tons)
log(Permit Price)	-352.9**	-485.4***	13.28	-27.20**
	(154.7)	(164.2)	(37.65)	(12.56)
ARPII	3,159***	3,201***	174.3***	140.4***
	(191.2)	(201.9)	(46.31)	(26.32)
Affected	-35.85	440.8**	139.4***	15.91
	(196.0)	(207.8)	(47.65)	(16.94)
				-
DID	-2,150***	-2,361***	-318.8***	132.7***
	(321.8)	(340.9)	(78.17)	(29.59)
priceDID	-206.7***	-226.9***	-20.56	10.10***
	(52.74)	(56.06)	(12.86)	(2.988)
log(Operating Capacity)	1,057***	1,064***	352.9***	7.436
	(69.89)	(74.19)	(17.01)	(5.192)
				-
Natural gas as primary fuel	-586.3***	-268.3*	-142.0***	118.9***
	(133.1)	(140.8)	(32.29)	(9.883)
Natural gas as secondary fuel	-423.1*	-62.16	-126.1**	216.6***
	(246.9)	(262.3)	(60.15)	(17.02)
Constant	-873.7	637.8	-1,031***	81.88
	(864.1)	(918.1)	(210.5)	(59.70)
Observations	2,072	2,082	2,082	976
R-squared	0.422	0.443	0.297	0.515

Standard errors in parentheses

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

Note: The table reports regression results from Model (1) in level forms; the coefficient estimations here have same signs as in the log-form results.

Table C. Regression Coefficient Estimation Results of Sulfur Factor Model in Level Forms

Sulfur Factor	(1) OLS	(2) OLS	(3) GLS	(4) GLS
Average Permit Price (dollars per ton)	2.64e-06	-3.90e-06*	2.64e-06	-3.90e-06*
	(1.96e-06)	(2.32e-06)	(1.94e-06)	(2.21e-06)
ARPII	0.000191***	1.74e-05	0.000191***	1.74e-05
	(3.95e-05)	(0.000191)	(3.91e-05)	(0.000182)
Natural gas as primary fuel	-0.00106***		-0.00106***	
	(5.06e-05)		(5.00e-05)	
Natural gas as secondary fuel	0.000390***		0.000390***	
	(9.17e-05)		(9.06e-05)	
Constant	0.000393***	0.000973***	0.000393***	0.000973***
	(6.95e-05)	(0.000133)	(6.87e-05)	(0.000127)
Including units burning natural gas	Y	N	Y	N
Observations	916	203	916	203
R-squared	0.491	0.884		
Number of identification			136	34

Standard errors in parentheses

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

Note: This table contains coefficient estimations for Model (3) in level forms; the first two columns are OLS estimation results and the last two columns are GLS estimation results; Column 2 and 4 shows results of analyzing on units not burning natural gas; the coefficient estimations here have same signs as results in log forms.