Incentives to Invest, Storable Permits, and Transaction Costs in Market-Based Environmental Regulation*

Yuta Toyama†
Northwestern University

Job Market Paper

January 14, 2018
Latest version is here:
http://sites.northwestern.edu/ytw959/

Abstract

Pollution reduction requires a costly investment in clean technology. The regulator often imposes a uniform standard on emissions intensity to induce abatement investments. However, such an approach may not be cost-effective in comparison with market-based regulations because firms are heterogeneous in their abatement costs. In this paper, I study the welfare effects of a cap-and-trade program. I develop an empirical framework that incorporates forward-looking behavior and transaction costs. In the presence of transaction costs in permit trading, investment patterns may depart from the first best outcome, violating the Coase theorem. Storable emissions permits allow firms to smooth costs over time, and abatement investment introduces dynamic incentives into compliance decisions. I apply the framework to study the first nine years (1995-2003) of the US Acid Rain Program. Using data on permit transactions and electricity production, I estimate the model and show that variable transaction costs are substantial. I use the estimated model to quantify the effect of a cap-and-trade program in comparison to a uniform standard, given a fixed level of aggregate emissions. I find that the total costs of reducing emissions under cap-and-trade are 16.6% lower. Although health and environmental damages from SO_2 emissions increase due to the change in the geographic distribution of emissions, the net benefit of the cap-and-trade is positive. I also examine the potential gains from trade in the absence of transaction costs. I find more dispersed patterns of investment and less banking of permits, both of which result in cost savings.

*I would like to thank Rob Porter and Mar Reguant for their guidance and encouragement throughout the project. I am grateful to Igal Hendel, Gaston Illanes, Koichiro Ito, Kei Kawai, Lola Segura, Max Tabord-Meehan, Yuta Takahashi, Nicky Vreugdenhil, Yasutora Watanabe, and seminar participants at Northwestern University, Waseda University, Kyoto Summer Workshop on Applied Economics 2017, Heartland Environmental and Resource Economics Workshop at Illinois 2017, and EMCON 2017 for helpful comments and discussions. All errors are mine.

†Department of Economics, Northwestern University. E-mail: yuta-toyama@u.northwestern.edu

1 Introduction

How to achieve environmental sustainability without compromising economic efficiency is a central question in policy debates and the economics literature. A traditional approach to pollution regulation is imposing a uniform standard on emissions intensity, although this approach may not be a cost-effective solution. Pollution abatement often requires costly investments in clean technology, and firms are heterogeneous in their costs of reducing pollution. Instead, economists have advocated market-based solutions to give firms an incentive to internalize negative externalities. One example is a cap-and-trade program where firms trade emissions permits to achieve a target level of aggregate emissions. When assessing the welfare effects of a cap-and-trade system, two critical elements exist: transaction costs and dynamic regulatory environment. Transaction costs associated with permit trading may prevent the first best reallocation of emissions permits, violating the Coase theorem. Dynamics influence how firms make investment decisions in the costly compliance technology. In this paper, I develop an empirical framework for a cap-and-trade program that incorporates firms' forward-looking behavior and transaction costs. I apply this framework to evaluate the welfare effects of the US Acid Rain Program, a federal cap-and-trade program designed to reduce sulfur dioxide emissions.

In a cap-and-trade program, firms face two key dynamic decisions: whether to invest in clean technology, and whether to store (bank) emissions permits. Investments in clean, but costly, technology are an important margin for reducing emissions. A cap-and-trade program typically spans a long time horizon, and the allocation of emissions permits changes over time. Firms, therefore, must take into account the change in the regulatory environment in their investment decisions. Moreover, firms can store (bank) emissions permits across periods. The storability of emissions permits under a permit banking system allows firms to smooth their costs over time. Although these dynamic factors are key elements of cap-and-trade programs, modeling these factors and estimating such a dynamic model is a challenging task. To deal with these complications, previous studies focused on a static decision problem on the steady state (See, e.g., Carlson et al. 2000, Fowlie 2010b, and Chan 2015). An innovation of my paper is to model explicitly dynamic aspects of a cap-and-trade program and bring it to the data, which allows me to evaluate comprehensively the evolution of a cap-and-trade program.

Another important focus of my paper is the role of transaction costs in the permit market. Transaction costs discourage firms from relying on the permit trading as a compliance strategy. Previous studies documented that many firms tend not to trade emissions permits, and instead comply with the regulation using their allocated permits (see, e.g., Jaraitė-Kažukauskė and Kažukauskas, 2015, for the EU Emissions Trading Scheme). In the absence of transaction costs, the Coase (1960) theorem implies that investment in abatement (i.e., reduction of emissions) should be efficient under cap-and-trade programs. In practice, transaction costs distort investment incentives, leading to the efficiency loss. My paper quantifies transaction costs in the permit market and their impacts on investment patterns and abatement costs.

I construct a dynamic equilibrium model of investment and a cap-and-trade program that captures these complex interactions in an equilibrium framework. The novelty of my model is that it includes dynamic investment decisions, permit banking, and transaction costs, each of which was studied separately in the previous literature, in a unified framework.[1] In my model, firms are price takers for emissions permits. They face various tradeoffs in their compliance decisions. Firms can comply with the regulation either by reducing emissions, which can be achieved by decreasing production or investing in clean technology, or by buying emissions permits. However, firms incur two type of transaction costs in the permit market: (1) a sunk cost associated with participation in the permit market and (2) variable costs that depend on the trading volume. These costs discourage firms from permit trading and affect investment patterns. The sequence of permit prices is determined by market clearing conditions, resulting in a dynamic competitive equilibrium.

I apply this framework to study the first nine years (1995-2003) of the US Acid Rain Program, a cap-and-trade program for sulfur dioxide (SO_2) emissions that targets the US electricity industry.[2] The goal of the Acid Rain Program is to reduce aggregate SO_2 emissions from generation facilities to half of their 1980 levels. The regulator distributed emissions permits to the existing generation facilities, and these facilities were required to hold sufficient permits to offset their emissions each year.[3] Regulated sources could choose how to comply with the regulation. For example, they could switch to a cleaner fuel, invest in abatement equipment, or obtain additional permits from the market.

An appealing feature of the Acid Rain Program is the availability of data. Various information is publicly available, including production data for power plants and trading data for emissions permits by electric utilities. Combining these two data sets allows me to identify the key parameters of my model, including transaction costs of permit trading and investment costs. My focus on the Acid Rain Program is also motivated by previous findings that the gains from permit trading might not have been fully realized (Carlson et al. 2000 and Keohane 2006 on Phase I, and Chan 2015 on Phase II). I revisit their findings by taking into account transaction costs of permit trading as a source of inefficiency.

I combined data on permit transactions and production information between 1990 and 2003. Identification of my model relies on optimality conditions regarding firms' decisions and detailed information on production and permit transactions. I use production data before and after the introduction of the cap-and-trade program to estimate the firm-level profit function from electricity production. The profit function implies the marginal profit from emissions across firms. Absent variable transaction costs, marginal profits should be equalized across firms, and equal to the permit price. Variable transaction costs are identified from how marginal profits vary with the

[1] The previous literature that studies various models of a cap-and-trade program includes a static trading model with transaction costs (Stavins, 1995), a theoretical model of permit banking (Rubin, 1996; Schennach, 2000), and a model of long-run abatement investment (Fowlie, 2010b).

[2] I choose 2003 as a terminal period of my analysis because of the announcement of the Clean Air Interstate Rule in December 2003, which had a major impact on the regulatory environment regarding SO_2 emissions. See section 2.2 for details.

[3] Emissions permits are called emissions "allowances" in the Acid Rain Program because the term "permit" has another meaning in US environmental law. Because "permit" is the standard terminology in the economics literature, I use the term "permit" in this paper.

trading volume. Firm-level participation in permit trading is employed to identify sunk costs of participation. Finally, I use the first-order-condition for investment to identify the marginal costs of investment.

I estimate the model parameters by simulated nonlinear least squares. The estimates imply that sunk participation costs are quite small, but variable transaction costs from permit trading are substantial. The median marginal transaction cost is estimated to be 98 USD, while permit prices range between 100 and 200 USD in my sample period.

I conduct a series of counterfactual simulations using the model estimates. First, I examine the impact of a cap-and-trade program in comparison to a uniform standard regulation under which all firms are required to have the same emissions rate. I find that the investment pattern under cap-and-trade differs significantly from the pattern under a uniform standard regulation, for a given aggregate level of emissions. Rather than following the uniform emissions rate imposed by the regulator, firms can optimally choose their levels of investment based on the costs and returns under a cap-and-trade program. The aggregate costs of abatement are 16.6% lower than under a uniform standard regulation.

I also discuss the implications for health and environmental damages. A potential concern of a cap-and-trade program is that it could lead to higher health and environmental damages in comparison to uniform standard regulations. Even though the aggregate level of emissions is fixed, the geographic distribution of emissions might differ from that under a uniform standard. Damages from SO_2 emissions depend on the location of emissions sources. health and environmental damages, therefore, can differ across regulatory regimes (see, e.g., Muller and Mendelsohn 2009, Fowlie et al. 2012, Fowlie and Muller 2013, and Chan et al. 2015). I use the data from Muller and Mendelsohn (2009) to calculate the damages under both a cap-and-trade program and a uniform standard. I find that the health and environmental damage does increase under a cap-and-trade program. But the net benefit of the cap-and-trade, calculated by the sum of abatement costs and health and environmental damages, is positive.

In a second counterfactual simulation, I examine potential gains from trade. I find that shutting down transaction costs, as might happen if there was a centralized trading platform, would lead to a more dispersed distribution of emissions and investment levels, reflecting more active trading of emissions permits. There is less permit banking in the absence of transaction costs. Transaction costs discourage firms from selling emissions permits, and firms prefer to accumulate these permits, lowering the allocative efficiency of emissions permits. In the absence of transaction costs, the total abatement costs decrease by around 37%, enough to offset increases in health and environmental damages. Thus, “unrealized” gains from trade are significant in my sample period.

My empirical framework can be applied to other market-based environmental policies, including water trading systems, the current Corporate Average Fuel Economy (CAFE) regulation, and the Renewable Portfolio Standard. A key feature of these regulations is the interaction between investment in clean technology and trading of environmental credits. For example, the recent CAFE standard regulation allows firms to trade CAFE credits with other firms for their compliance. This trading scheme is an alternative to improving fuel efficiencies for a manufacturers' fleet by making a costly investment. Under the Renewable Portfolio Standard, electricity utilities can either invest in renewable technologies or obtain the Renewable Energy Certificates from the market to comply with the regulation.

1.1 Related Literature

My paper is related to three strands of literature: (i) the empirical literature on dynamic investment behavior, (ii) the empirical literature on cap-and-trade regulations, and (iii) the evaluation of the Acid Rain Program.

Understanding firms' investment behavior and its welfare consequences is a central theme in the industrial organization literature. Previous work includes Ericson and Pakes (1995), Bajari et al. (2007), Ryan (2012), Collard-Wexler (2013), and Kalouptsidi (2014) in an oligopolistic setting, and Rust (1987), Aguirregabiria and Mira (2002), and Kellogg (2014) in a competitive setting.

A novel feature of my paper is that investment in technology is substitutable with trading of emissions permits. To comply with the cap-and-trade regulation, a firm can either make an investment and reduce emissions, or purchase emissions permits from the permit market, where a firm faces transaction costs. Investment decisions also interact with storability of inputs, namely banking of emissions permits in my model. For example, if tighter future regulatory intensity is announced at the inception of the regulation, permit banking induces investment in early periods. A similar market structure can be found in other settings, including the CAFE credit trading program and the green certificate trading in the Renewable Portfolio Standard.

My paper also contributes to the empirical literature on cap-and-trade programs. Much of the literature test qualitative predictions of models of permit trading. A few recent papers take a structural approach to measure the welfare implications of permit trading.[4] In the context of NOx regulation, Fowlie (2010b) constructs a model of abatement choice to study the effect of rate-of-return regulation on permit trading. Fowlie et al. (2014) construct and estimate a model of dynamic investment and entry/exit game to discuss the implications of hypothetical market-based environmental policies in the US cement industry.[5] Abito (2014) quantifies the impact of rate-of-return regulation on the efficiency of SO_2 emissions regulation by estimating a multi-product cost function, though he does not explicitly consider permit trading.

A distinctive feature of my paper is to model trading behavior in the permit market and banking of emissions permits.[6] The previous papers all assume frictionless permit markets in which cap-and-trade is equivalent to imposing a Pigouvian tax. My model captures emissions permit trading when permits may be banked and transaction costs exist. It can quantify transaction costs and

[4] The literature has examined the independence of outcomes from the initial allocation (Reguant and Ellerman, 2008 and Fowlie and Perloff, 2013) and the internalization of emissions costs (Kolstad and Wolak, 2008, Fowlie, 2010a, and Fabra and Reguant, 2014).

[5] Dardati (2014) also studies how an allocation scheme for closing plants affects entry/exit decisions, using the calibrated model of industry dynamics in the context of the Acid Rain Program.

[6] Cantillon and Slechten (2015) might be the closest to my paper. They study participation decisions and price formation for CO_2 emissions permits using trading data in the EU-ETS scheme.

their impact. My framework can also be used to study how the regulatory design of permit trading, such as the availability of permit banking and alternative allocation rules for emissions permits, affects firms' abatement decisions.

Finally, my paper provides new insights for the evaluation of the Acid Rain Program. One approach in the literature is to calculate cost saving due to permit trading by estimating a cost function and a discrete choice model for abatement choice (see, e.g., Ellerman et al., 2000, Carlson et al., 2000, Keohane, 2006, and Chan, 2015). Researchers found that adopting a permit trading program led to significant cost savings compared to traditional command-and-control approaches, though the actual cost did not reach the least-cost solution. Another approach is to focus on aggregate variables to discuss the efficiency of the permit market (Joskow et al. 1998, Ellerman and Montero 2007, and Helfand et al. 2006).[7]

My paper complements this literature by providing an empirical model that incorporates firms' decisions on abatement, permit trading, and permit banking in an equilibrium framework. My model allows me to evaluate the role of permit banking and transaction costs. I decompose the effects of the Acid Rain Program into the effects of permit trading and permit banking. Previous studies often note the importance of permit banking as a source of cost efficiency. My paper is the first to quantify the gains. Moreover, I quantify the potential gains from trade which could be achieved in the absence of transaction costs. These simulation analyses require an equilibrium model of the cap-and-trade program. I also use transaction data for emissions permits to estimate the model, in contrast to most of the literature.

2 Empirical Setting and Descriptive Analysis

2.1 The Acid Rain Program

Fossil-fuel electricity plants, especially coal plants, produce sulfur dioxide (SO_2) emissions as a byproduct of electricity generation. SO_2 is known to have detrimental effects on human health and the environment. Although the federal government introduced command-and-control-type regulations with the Clean Air Act Amendments of 1970, such regulations have not been effective in reducing SO_2 emissions.[8] The failure of the previous regulations led to the introduction of the Acid Rain Program (ARP), a cap-and-trade program, in 1995.

The target of the regulation is electricity generating units (EGUs) that use fossil fuels and have an output capacity greater than 25 megawatts. The regulation was implemented in two phases. In Phase I (1995-1999), a subset of eligible EGUs are under the regulation. These units include 263 EGUs named the "Table 1" group, which were especially dirty and old before the regulation, and

[7] Joskow et al. (1998) finds that prices in the spot market and the EPA auction are close and concludes that "a relatively efficient private market" had developed by mid-1994. Ellerman and Montero (2007) argues for the efficient market of permits by comparing the actual and theoretically predicted volume of aggregate banking. Helfand et al. (2006) uses the monthly permit prices during 1994 to 2003 to test whether the price path follows the Hotelling r-percent rule for intertemporal arbitrage. They reject the Hotelling rule, which is a suggestive evidence of inefficiency of the market.

[8] Ellerman et al. (2000) provide a brief history of the regulation on SO_2 emissions.

an additional 182 EGUs from the Non-Table 1 group as substitution or compensating units. In Phase II (begun in 2000), all eligible EGUs are mandated to comply with the regulation.

The ARP aims to reduce SO_2 emissions from generation facilities to half of their 1980 levels, which determines the total number of emissions permits for each year. Most of the emissions permits are allocated for free to incumbent units (grandfathering scheme). The EPA adopts the rule that determines the unit-level allocation of emissions permits based on the characteristics of a unit.[9] The allocation is primarily determined by the product of average heat inputs during 1985-1987 and the target emissions rate for each Phase (2.5 pounds per 1 million British thermal unit (lb/MMbtu) in Phase I and 1.2 lb/MMBtu in Phase II). Some units also obtain additional allocation of permits based on technical and political considerations (Joskow and Schmalensee, 1998).

SO_2 permits are tradable goods. Firms can sell or buy permits with other firms, including financial companies or brokers that do not own any generating units and thus are not required to comply with the regulation. Although the EPA also holds an annual auction to distribute around 2.7% of the yearly allocation, a centralized trading exchange does not exist. Bilateral trading, which is often mediated by brokers, is the primary way to trade emissions permits with other participants.

Operation of each regulated unit, especially emissions levels of SO_2, is recorded through the Continuous Emissions Monitoring System. At the end of the calendar year, the annual level of SO_2 emissions is finalized, and each regulated unit is required to surrender emissions permits within a grace period of 60 days. The remaining permits are carried over to the next year, which is called banking of emissions permits.[10] As I discuss in section 2.3, regulated firms had a significant amount of banked permits in Phase I when the grandfathered allocation was more generous than in Phase II.

Although emissions permits were grandfathered to the existing units, most of them still needed to decrease their emissions from their business-as-usual level to comply with the regulation. The regulated units were able to reduce emissions by either lowering utilization (output) or emissions per output (emissions rate). The latter option of reducing emissions rates was the primary channel of abatement, which I will explain in section 2.3 in detail.

2.2 Data

In this paper, I focus on the period from 1995 to 2003. Although the ARP continued after 2004, the proposal of the Clean Air Interstate Rule, announced in December 2003, had a large impact on the regulated firms' expectation over the future regulatory environment. The proposed regulation aimed to strengthen the stringency of the SO_2 regulation from 2010 in the framework of the ARP. After the announcement, the permit price started to rise dramatically, primarily because the value of emissions permits issued before 2010 would be higher than those issued after 2010 in the proposed regulation. Firms also started to invest in scrubbers in anticipation of more strict intensity of the

[9] See U.S. Environmental Protection Agency (1993a,b) for the details.

[10] If an affected unit does not hold sufficient permits to offset the emissions at the end of the compliance deadline, unit operators are required to pay the penalty of $2000 per SO_2 ton. However, compliance was nearly 100% during the period of my analysis.

proposed regulation.[11] Thus, I do not include the data after 2004 and rather focus on the periods when the regulatory environment regarding SO_2 emissions was stable.

The data I use are a combination of transaction data for emissions permits and various data on electricity production. The data on permit transactions are from the Allowance Tracking System (ATS) operated by the EPA. The latter data are a compilation of various databases from the EPA and the US Energy Information Administration (EIA). I explain these two type of data in turn.

First, the EPA uses the ATS to manage permit allocation and track private transactions and surrender of permits for compliance, and makes the data public. Each transaction record in the tracking system contains the account name of a transferor, a transferee, vintage of permits, quantity of transferred permits, and the confirmation date of the transaction.[12] I constructed the transaction data at the firm and year level from the database. Specifically, I aggregated the account-level information into the firm-level information by using ownership information constructed from various sources including EGrid database and EIA-860. The final data set includes (1) permit holding at the beginning of the year, (2) grandfathered allocation, (3) volume of permit transaction (net purchase of emissions permits), and (4) banking volume. Note that I only used information on transactions of permits whose vintage is current or old.

The ATS does not collect any information on transaction prices. I instead collected the market-price index of SO_2 permits provided by Cantor Fitzgerald, one of the biggest brokers in SO_2 permit markets. The frequency of the price data is monthly. I explain the details of the price data in section 2.3.5.

The second piece of my data set is production information of electricity companies. I combined multiple databases to construct the data set. These databases include EPA data as well as EIA survey data. First, the EPA makes public unit-level operation data of generating units collected by the CEMS. The CEMS data include gross generation (in MWh), heat inputs (in MMBtu), and SO_2 emissions. In addition, the EIA conducts various surveys on operation of power plants. Specifically, the Form EIA-767 "Steam-Electric Plant Operation and Design Report" gives me information on fuel usage (sulfur content, ash content, heat inputs), net generation, and generation capacity at the unit and month level. Also, the Form FERC No. 423 (EIA-423) "Monthly Report of Cost and Quality of Fuels for Electric Plants" provides plant- and month-level information on fuel procurement, including fuel type, sulfur contents, heat contents, and purchase costs.

2.3 Descriptive Analysis

I now provide a descriptive analysis on the data set I constructed. I focus on various aspects of the ARP including banking of emissions permits, abatement decisions of regulated sources, and market of emissions permits. These descriptive findings motivate the modeling approach I introduce in

[11]See Schmalensee and Stavins (2013) for a detailed review on how the regulatory environment regarding SO_2 emissions has been changing since 2004.

[12]The confirmation date must lag behind the actual transaction date to some extent, although the prompt recording of private trading was considered the rule rather than the exception according to the EPA staff and industry experts. See Joskow et al. (1998) for details.

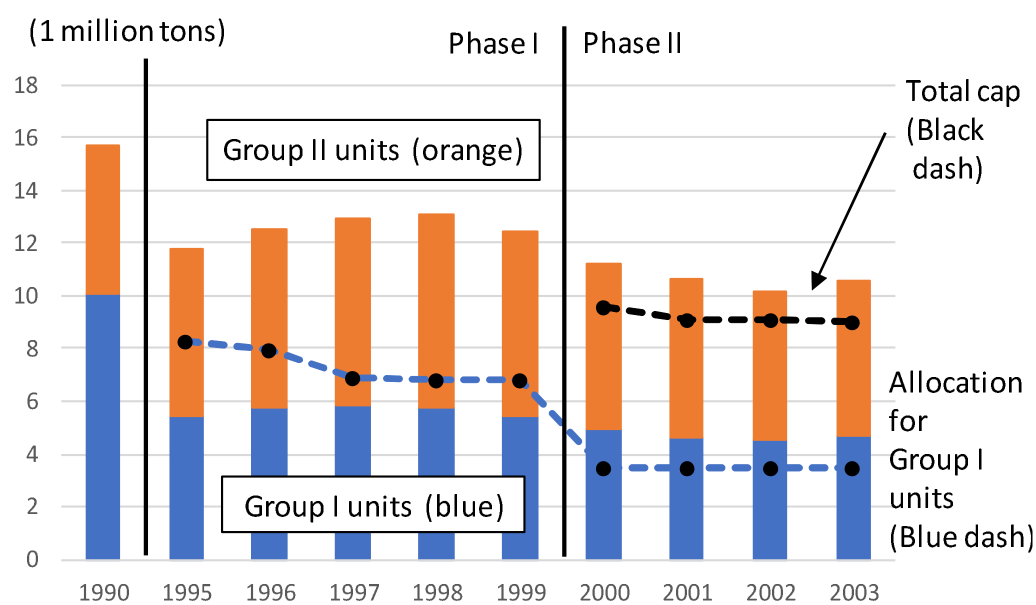
section 3.

2.3.1 Banking of Emissions Permits

Figure 1 shows the aggregate SO_2 emissions level and emissions caps under the ARP from 1990 to 2003. The bars show emissions levels each year, and the dashed lines show the emissions cap. As I mentioned in section 2, the timing of the regulation was different across electricity generating units. I denote those units that are regulated beginning 1995 as group I units and those regulated since 2000 as Group II units. The blue bar in the figure corresponds to emissions from Group I units, and the orange bar corresponds to those from group II units. The blue dashed line shows the allocation for Group I units, and the black dashed line from 2000 shows the total cap of emissions including both Group I and II units.

The figure shows that Group I units reduced their emissions almost by half compared to their 1980 level once Phase I started in 1995. While both Group I and II units reduced emissions further in 2000, the first year of Phase II, Group I units did not reduce emissions as much as in 1995. Emissions before 1999 were significantly lower than the emissions cap, though the aggregate emissions exceed the allocation of emissions permits after 2000. These observations imply that Group I units saved their permits in Phase I and then started to use them after 2000 for the purpose of compliance.

Figure 1: Aggregate Volume of SO_2 Emissions and Caps (1990 - 2003)



Notes: The blue (orange) bar corresponds to emissions from Group I (Group II) sources. The blue dashed line shows permit allocation for Group I units, and the black dashed line from 2000 shows the total cap including allocation for both Group I and II units.

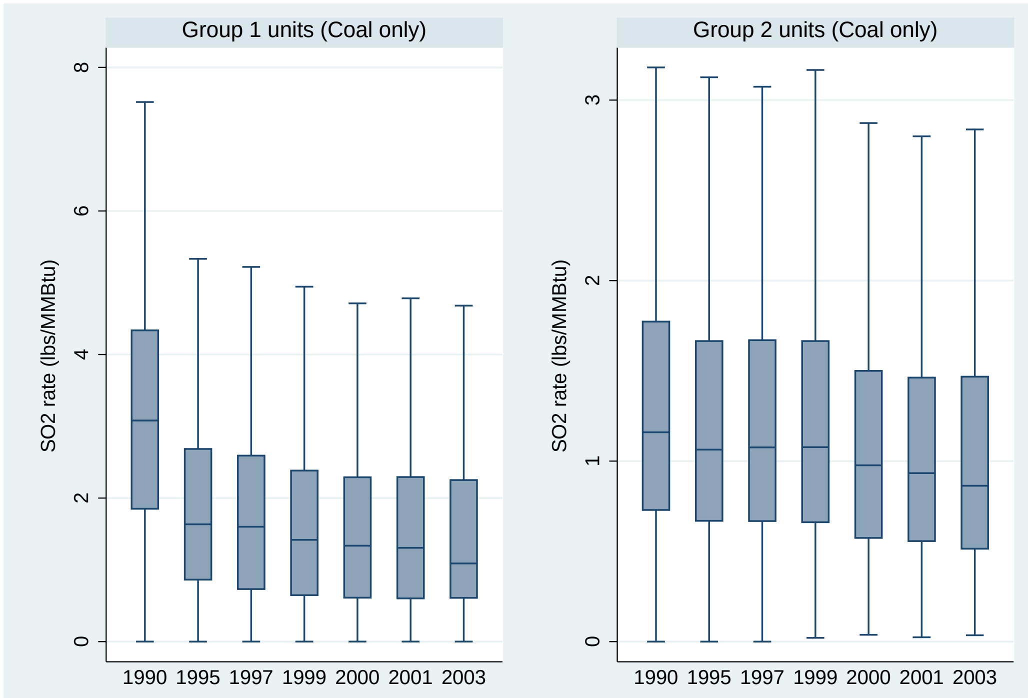
2.3.2 Abatement Strategy for Coal units

Emissions from electricity generation can be reduced either by (1) reducing the emissions rate (i.e., emissions per output) or (2) reducing output (lower utilization of generating units). In this subsection, I explain the former abatement strategy for generating units whose primary fuel type is coal. Although the target of the ARP includes all types of fossil fuel units (coal, gas, and oil), SO_2 emissions from gas and oil units are relatively small and no room remains for lowering the emissions rate of these units.

Two common options are available to reduce the emissions rate of coal units. The first option is called fuel switching. An operator of coal units can switch the type of coal from dirty (e.g., high-sulfur bituminous coal) to cleaner (e.g., subbituminous coal or low-sulfur bituminous coal). The fuel costs of cleaner coals are higher than the fuel costs of dirty coals. Also, switching fuel types requires retrofitting the boiler to make it compatible with the new type of coal, which incurs fixed costs. Another abatement option is installing flue-gas desulfurization equipment (a scrubber). This equipment is installed at the stack of generation units and eliminates more than 80% of SO_2 emissions. This option, however, incurs large investment costs as well as a long lead time (2 to 3 years on average).

Figure 2 shows the distribution of unit-level SO_2 emissions rates (measured in pounds per MMBtu) for each group in selected years. The left panel shows the distribution for group I sources. The emissions rates of these sources decreased between 1990 and 1995, the beginning of Phase I. The emissions rates stayed almost constant within Phase I, and it decreased further in 1999, which anticipates the beginning of Phase II. For generating units in group II, their emissions rates did not change until 1999, and then decreased in 2000, the first year of the cap-and-trade program for these units. These observations imply that firms adjusted their emissions rates at the beginning of each phase, but emissions rates remain almost constant within the phase. This observation motivates my model of abatement investment in the structural model I introduce in section 3.

Figure 2: Distribution of Unit-level SO_2 Emissions Rate



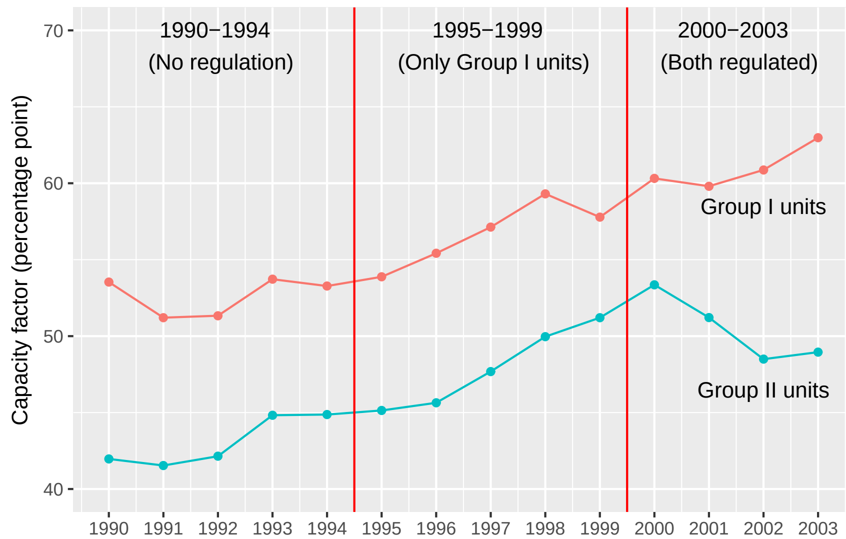
Note: The box shows the interquartile range of the distribution. The two lines correspond to the upper and lower adjacent values of the distribution.

2.3.3 Effects of a cap-and-trade on Output

I now examine whether firms reduced outputs in response to the regulation, which is another margin of emissions abatement. Here, I focus on the intensive margin of operation and treat entry/exit as given. Although retirement of coal units could be a potential option for emissions abatement, the data shows that this margin is small. Among the 263 EGUs in the "Table 1" group, only seven units retired before 1995, and two additional units retired before 2003. Regarding other coal units, around 6% of EGUs retired between 1990 and 2003.

To estimate the effects of the ARP on production output, I exploit the variation of the timing of the regulation across units in a difference-in-differences (DID) framework. Figure 3 shows the trend of the capacity factor, defined by the ratio of net generation (output) to generation capacity, over time. I calculate the mean of the monthly level capacity factor in each year for two groups: those that are regulated from 1995 (Group I units) and those that are regulated from 2000 (Group II units). The figure shows that these two groups have a similar trend in their capacity factor from 1990 to 1994, which supports the parallel-trend assumption in the DID framework.

Figure 3: Trend of Capacity Factor of Group I and Group II units



The regression equation I estimate is given by

$$cf_{jm} = \alpha_1 \text{GroupI}_j \cdot 1\{t \geq 1995\}_t + \alpha_2 \text{GroupII}_j \cdot 1\{t \geq 2000\}_t + \gamma X_{jm} + u_j + u_m + u_{jm},$$

where cf_{jm} is capacity factor of unit j in month m. The capacity factor is defined by $cf_{jm} = q_{jm}/k_j$, where q_j is net-generation and k_j is nameplate capacity. GroupI and GroupII are the dummy variables for each group. X_{jm} includes control variables such as fuel costs. Unit and time fixed effects are captured by u_j and u_m.

Regression results are shown in Table 1. I find that the introduction of the ARP decreased the capacity factor by 1 to 2.5 percentage points, which is statistically significant. This finding is consistent with the idea that introducing a cap-and-trade program increases marginal costs of production, because firms are facing opportunity costs of emissions under a cap-and-trade program. The increase in marginal costs thus decreases outputs of generating units under a cap-and-trade regulation. Although the effects are statistically significant, the economic significance of the effects seems to be limited. Because the mean of the capacity factor is within the range of 40-60 percentage points in my sample, electricity generation decreased by around 2%-6% due to the introduction of a cap-and-trade program. This magnitude is not as much as the decrease in emissions over time, as shown in section 2.3.1. Combined with the findings from the previous subsections, this regression analysis indicates that the abatement of SO_2 emissions was achieved primarily through the adjustment of emissions rates.

Table 1: Reduced-Form Model of Capacity Factor

	Dependent variable:			
	Capacity factor in pct-point			
	(1)	(2)	(3)	(4)
Treatment (Group I units)	-0.656	-2.115^{***}	-1.141^{**}	-2.517^{***}
	(0.575)	(0.668)	(0.562)	(0.693)
Treatment (Group II units)	-4.002^{***}	-2.615^{***}	-2.363^{***}	-1.082^{*}
	(0.548)	(0.592)	(0.564)	(0.619)
log(fuel costs)			-11.214^{***}	-11.168^{***}
			(0.479)	(0.479)
log(electricity demand)	40.937^{***}	40.986^{***}	42.479^{***}	42.494^{***}
	(1.194)	(1.195)	(1.229)	(1.231)
Group-trend	No	Yes	No	Yes
Observations	306,727	306,727	252,223	252,223
Adjusted R^2	0.635	0.635	0.600	0.600

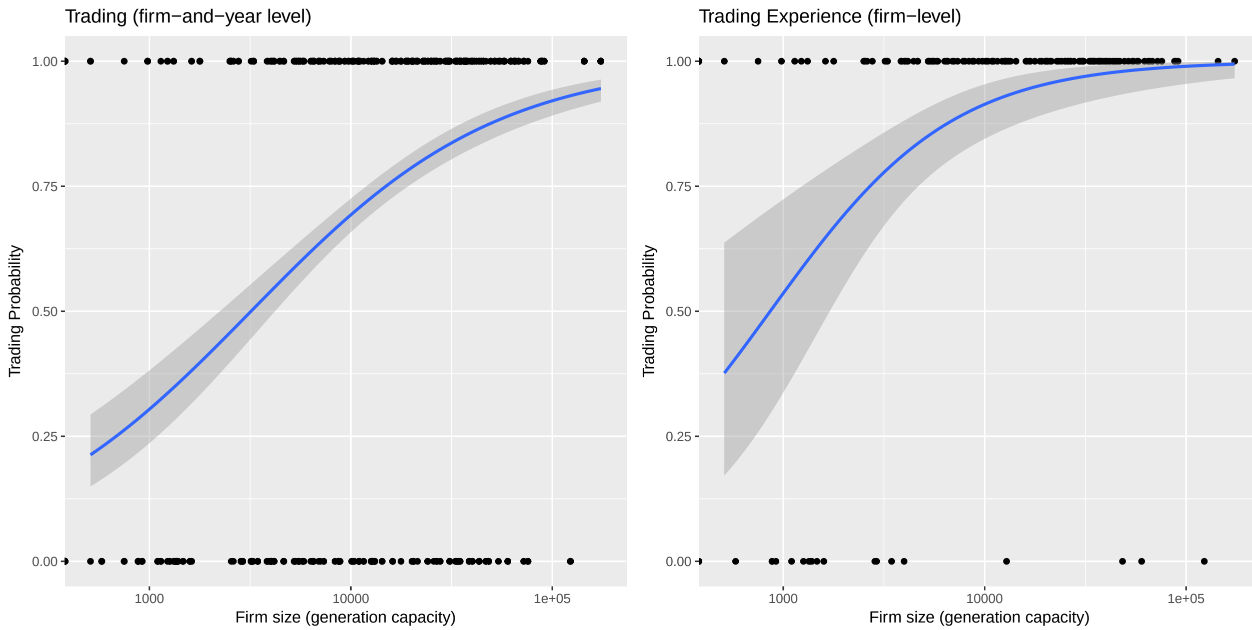
Notes: Unit-level dummies, year dummies, and month-of-year dummies are included. Standard errors are clustered at the unit level. $*p < 0.1; **p < 0.05; ***p < 0.01$

2.3.4 Firm-level Trading Information

I now explain how firms behaved in the market of emissions permits. Figure 4 shows the correlation between trading decisions and firm size, measured by the sum of nameplate capacity of units under the ARP. The left panel shows the unconditional probability of market transaction at the firm-year level, and the right panel shows trading experience in the sample period at the firm level.

The left panel shows that firms did not necessarily trade every year. The unconditional probability of conducting permit trading was 72%. The trading probability was positively correlated with firm size. This observation is also found in the context of the EU-ETS scheme (see, e.g., Jaraitė-Kažukauskė and Kažukauskas, 2015). Although this finding can be interpreted as suggestive evidence for the presence of fixed costs of transaction, firms do not need to conduct a transaction in every period, due to the storability of emissions permits. In the right panel, I show firm-level experience of market trading during the sample period. 86% of firms had at least one experience of trading in the sample period, although some firms, most of which are small, did not trade at all.

Figure 4: Trading Pattern at Firm Level



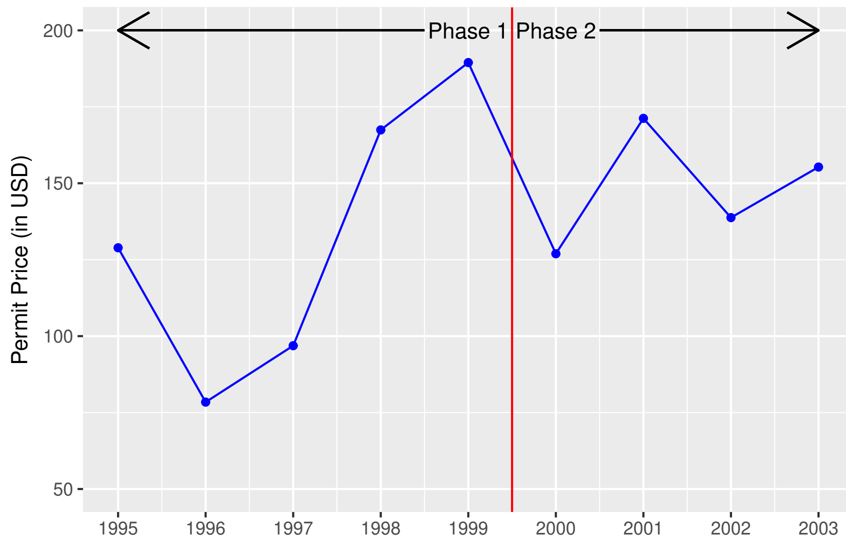
2.3.5 Price Data from a Broker

As I discussed in section 2.1, there is no centralized trading exchange for emissions permits under the Acid Rain Program. Although regulated firms need to have bilateral trades with other firms, brokers act as an intermediary for those transactions. Brokers also provide information about permit prices. Figure 5 shows price information provided by Cantor Fitzgerald, a broker in this market. I use the monthly SO_2 price index as a price measure in this paper. Cantor Fitzgerald constructs this index based on various trading information including the allowance bids (to buy), the allowance offers (to sell), and the actual trade prices. The company also posts this price information on the website in every month. I aggregate the monthly price index by taking the median for each year. Note that the price is also normalized to the level of 2000 by the Producer Price Index.

The price at the beginning was around 150 USD, and fell below 100 USD in 1996 and 1997. It increased to 200 USD in 1999, and fluctuated in the range of 120-200 USD after 2000. The figure suggests that the market price reflects the availability of banking. In the absence of permit banking, I would expect to see a spike in the permit price between Phase I and II, because the target emissions rate in Phase II is much stricter than in Phase I. Instead, the permit price has been gradually increasing over time, though it is volatile to some extent. [13]

[13] A key theoretical prediction regarding permit prices is the Hotelling rule: permit prices should increase with the risk-free interest rate if the market is efficient and there are no transaction costs. Helfand et al. (2006) test the Hotelling rule by using the monthly prices of emissions permits in the same period, and reject the rule after controlling for structural changes and market shocks.

Figure 5: Price of Emissions Permits by a Broker



Note: Price is normalized to January 2000 by the Producer Price Index.

3 Model of Cap-and-Trade Program

3.1 Overview of the Model

This section introduces a structural model of the cap-and-trade program. My model is a discrete- and finite-horizon model indexed by $t = 1995, \ldots, 2003 (\equiv T)$, and each discrete decision period corresponds to one compliance year. Firms have the common discount factor of β.

The overview of the model is summarized in Figure 6. The model has two building blocks: (i) investment in abatement options at the beginning of each phase (1995 and 2000), and (ii) decisions on production, trading, and banking in each year. At the beginning of each phase (1995 and 2000), firms make an investment decision on the abatement option and determine the emissions rate (R_i^1, R_i^2). The emissions rates are assumed to be fixed within each phase. This assumption reflects the observation from section 2.3 that the emissions rate changes at the beginning of each phase and stays constant within the phase.

Given the emissions rate, a firm makes decisions on production, permit trading, and banking. The timeline of each period is as follows:

1. Firm i holds carry-over permits h_{it} and receives grandfathered permits a_{it}.

2. Participation decision: Denote firm i's experience of market trading by I_{it}; i.e., $I_{it} = 1$ if a firm has experience in market trading, and 0 otherwise. If $I_{it} = 0$, a firm can pay the one-time sunk cost F_{it} to participate.

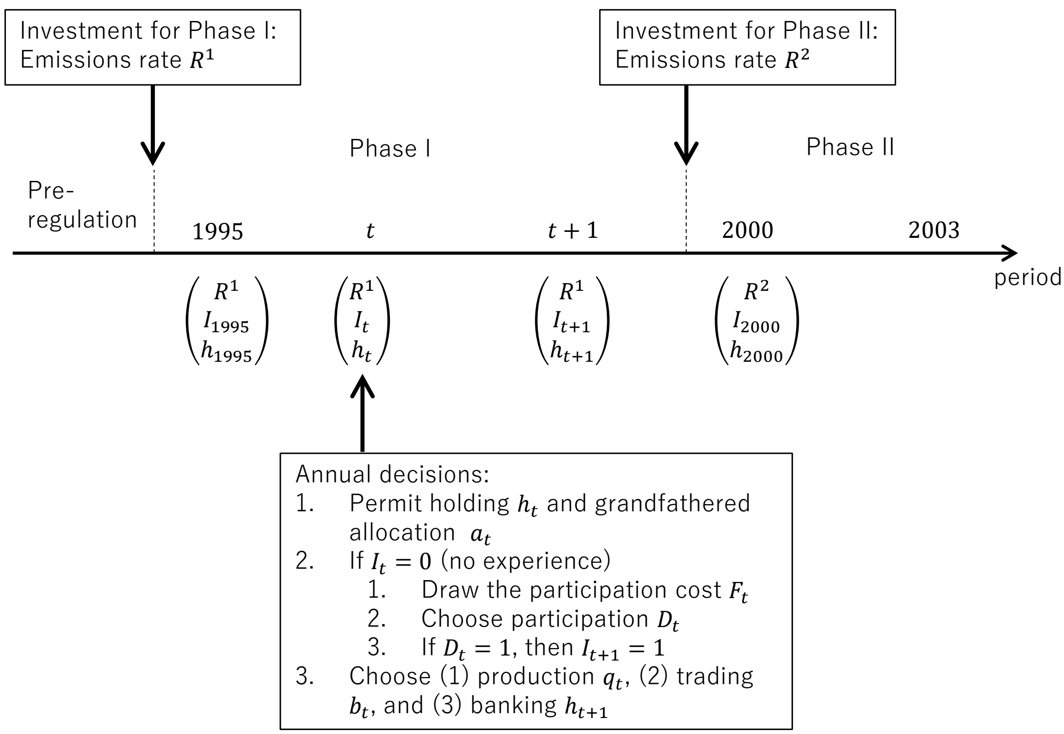
3. A firm chooses (i) production quantity of each generating unit $\{q_{jt}\}_j$, (2) net volume of trading b_{it} if a firm already participated in the market, and (3) banking of permits $h_{i,t+1}$.

4. A firm obtains profits from electricity generation and pays the costs of permits (or obtains the revenue from selling permits).

5. Move to the next period with the holding $h_{i,t+1}$.

I now turn to explain each component of the structural model.

Figure 6: Model Overview



3.2 Electricity Production and SO_2 Emissions

Firms are earning profits from electricity production in a competitive electricity market. Firm i holds J_{it} units of the regulated sources and chooses the production quantity of generating units $\{q_{jt}\}_j$. The profit is given by

$$\pi_{it}\left(\{q_{jt}\}_j\right) = \sum_{j \in J_{it}} \left\{ \left(\tau_{st}^{elec} - c_{jt}^{fuel}\right) \cdot q_{jt} - g(q_{jt}, k_j) \right\},$$

where τ_{st}^{elec} is the electricity price in state s where unit j is located, and c_{jt}^{fuel} is the unit-specific fuel costs of production. Fuel costs account for around 75% of total operating expenses (see EIA, 2012).

$g(q_{jt}, k_j)$ is the convex cost of production. This term captures the increasing costs of operation near the capacity constraint (see, e.g, Ryan, 2012).

Electricity production is associated with SO_2 emissions. Firm-level emissions are given by

$$e_{it}\left(\{q_{jt}, \rho_{jt}\}_j\right) = \sum_{j \in J_{it}} \rho_{jt} q_{jt}, \tag{3.1}$$

where ρ_{jt} is the unit-level SO_2 emissions rate per production. I assume that ρ_{jt} is given by

$$\rho_{jt} = \begin{cases} HR_j \cdot R_i^1 & if\ j = coal\ \&\ t \in [1995, 1999] \\ HR_j \cdot R_i^2 & if\ j = coal\ \&\ t \in [2000, 2003] \\ HR_j \cdot R_{j,t}^{gas} & if\ j = gas\ or\ oil. \end{cases},$$

The unit-specific heat rate HR_j is an inverse of the production efficiency measure. HR_j represents how much fuel (in MMBtu) is needed to produce 1 MWh of electricity. R_i^1 and R_i^2 are the firm- and phase-specific SO_2 emissions rate, the emissions level per 1 MMBtu units of fuel. These emissions rates are endogeneously determined by the investment decisions at the beginning of each phase. I treat gas and oil units separately from coal units because these units have already low SO_2 emissions rates.

Note that the profit $\pi_{it}(\cdot)$ does not include the costs associated with emissions. Firms should take into account the cost of using emissions permits in their production decisions. As I show in section 3.4, the optimal decision on production quantity reflects emissions costs as well as the output price τ_{st}^{elec} and the fuel cost c_{jt}^{fuel}. The profit function $\pi(\cdot)$ is interpreted as the gross profit from electricity production that excludes costs associated with the permit trading.

3.3 Structure of Permit Trading and Transaction Costs

The role of cap-and-trade regulation is to penalize emissions from production activity and incentivize firms to reduce emissions. This subsection introduces the regulation into my model.

Each firm is allocated the grandfathered allocation a_{it} in each period. Because the allocation plan was announced before the regulation, the sequence of $\{a_{it}\}_t$ is exogenous in the model. The firm also holds the emissions permits that are carried over from the previous period, denoted by h_{it}. A firm decides emissions level e_{it}, which is determined by production quantity $\{q_{jt}\}$ as given by equation (3.1), net purchase volume b_{it}, and banking volume $h_{i,t+1}$. b_{it} is positive (or negative) if firm i is a buyer (or a seller), implying that she is buying (or selling) $|b_{it}|$ units of permits.

The transition of permit holding is given by

$$e_{it} + h_{i,t+1} = a_{it} + h_{it} + b_{it}, \tag{3.2}$$

$$h_{i,t+1} \geq 0. \tag{3.3}$$

Note that equation (3.3) is the non-negativity constraint of banking and excludes the possibility of borrowing of permits from future grandfathered permits. I assume that firms achieve perfect compliance in my model. This is based on the fact that the compliance rate under this regulation is nearly perfect.

I model the permit market as a competitive market with transaction costs. The Acid Rain Program was a federal-wide program where many electric utilities, as well as financial companies, were participating. Exercising market power in the permit market was limited. [14] The presence of transaction costs reflects the fact that the vast majority of permit transactions were bilateral because there was virtually no centralized exchange for emissions permits. Incorporating bilateral trading of emissions permits into my model, however, is quite difficult because emissions permits are divisible objects and my model also features dynamic investment in clean technology and permit banking. I thus capture the nature of permit market by introducing transaction costs in a reduced form way.

Firms are price-takers in the permit market and face the market price P_t. In addition, they have to pay two types of transaction costs (see, e.g., Stavins, 1995). First, when a firm trades for the first time, it has to pay a sunk cost of participation F_{it}. This cost is motivated by the observation that some firms did not participate in the permit trading. An interpretation of F_{it} includes the costs associated with setting up a trading desk at the company and hiring a financial-trading expert. I specify F_{it} as the sum of a fixed cost F and an idiosyncratic cost ϵ_{it} given by

$$F_{it} = F + \epsilon_{it}, \epsilon_{it} \sim G(\cdot; \sigma_F),$$

where $G(\cdot; \sigma_F)$ is the cumulative distribution function of type I extreme value distribution.

Second, firms have to pay variable transaction costs associated with net purchase of permits b_{it}. This cost is given by

$$TC(|b_{it}|),$$

where $TC(\cdot)$ is a differentiable and strictly convex function. Variable transaction costs include brokerage commissions and bid-ask spreads. The convex nature of the cost function also captures the difficulty of large-scale transactions of emissions permits. Suppose that a firm wants to buy a certain amount of permits, but its trading partner cannot meet the demand. In such a case, a firm has to find another trading partner to buy more permits, and hence incurs a costly search process in a bilateral market. Convex transaction costs are employed in the theoretical literature in finance (e.g., Gârleanu and Pedersen, 2013, and Dávila and Parlatore, 2017) and also motivated by empirical findings (see, e.g., Breen et al., 2002, Lillo et al., 2003, and Robert et al., 2012).

In summary, the compliance costs (or revenue) from trading b_{it} units of permits is given by

$$P_t b_{it} + TC(|b_{it}|).$$

[14] Liski and Montero (2011) examined how the four biggest electric utilities (in terms of initial allocation) trades in the permit market. They found that their behavior is not consistent with the model of market power in a storable commodity market.

Introducing Fringe Firms The sample I use does not cover all firms participating in permit trading. For example, financial companies or brokers do not have any generation facilities, thus I do not include them. Also some electricity companies are excluded from the sample in the process of data cleaning. I call them fringe firms in permit trading. To deal with the presence of fringe firms, I introduce the demand function of firms outside my sample. I denote the total net purchase from fringe firms by $\bar{B}_t^{fringe}(P_t)$. I explain a specification and estimation approach of the fringe demand function in section 4.3.

3.4 Optimal choices on production, trading, and banking

I now consider the optimization problems in year t. A firm makes both discrete (participation) and continuous decisions on production, trading, and banking. I first explain the decision problems conditional on the status of trading participation. These problems characterize the values from participation and non-participation, which determines the optimal participation decision.

Let V_{it}^1 and V_{it}^0 be the optimal values when a firm participates in trading ("trader") and does not ("non-trader"). The Bellman equation for the "trader" is given by

$$\begin{aligned} V_{it}^1(h_{it}, R_{it}) = \max_{\{q_{jt}\}_{j\in J_i}, b_{it}, h_{i,t+1}} \quad & \pi_{it}\left(\{q_{jt}\}_j\right) - (P_t b_{it} + TC(b_{it})) + \beta EV_{i,t+1}(h_{i,t+1}, 1, R_{i,t+1}) \\ \text{s.t.} \quad & e_{it}\left(\{q_{jt}, \rho_{jt}\}_j\right) + h_{i,t+1} = a_{it} + h_{it} + b_{it}, \\ & h_{i,t+1} \geq 0. \end{aligned} \tag{3.4}$$

$EV_{it}(h_{it}, I_{it}, R_{it})$ denotes the ex-ante value function for firm i in period t when the firm holds h_{it} units of emissions permits, the trading experience is I_{it}, and the emissions rate is R_{it}. Recall that $I_{it} = 1$ if firm i already participated in the market trading by paying the participation cost.

When a firm is a non-trader, it does not choose the trading volume b_{it} by definition. The Bellman equation in this case is

$$\begin{aligned} V_{it}^0(h_{it}, R_{it}) = \max_{\{q_{jt}\}_{j\in J_i}, h_{i,t+1}} \quad & \pi_{it}\left(\{q_{jt}\}_j\right) + \beta EV_{i,t+1}(h_{i,t+1}, 0, R_{i,t+1}) \\ \text{s.t.} \quad & e_{it}\left(\{q_{jt}, \rho_{jt}\}_j\right) + h_{i,t+1} = a_{it} + h_{it}, \\ & h_{i,t+1} \geq 0. \end{aligned} \tag{3.5}$$

Note that the value functions $\{V_{it}^0(\cdot), V_{it}^1(\cdot)\}$ are indexed by t, which is meant to include all state variables except for h_{it}, I_{it}, R_{it}, and ϵ_{it}. I assume perfect foresight over the state variable in the next period except for the shock to the participation cost ϵ_{it}.

The optimality conditions for the traders are given by

$$P_t^{elec} - c_{jt}^{fuel} - g'(q_{jt}) - \lambda_{it}\rho_{jt} = 0 \tag{3.6}$$

$$\lambda_{it} = P_t + TC'(b_{it}) \tag{3.7}$$

$$\lambda_{it} = \beta\frac{dEV_{i,t+1}(h_{i,t+1}, I_{i,t+1}, R_{i,t+1})}{dh_{i,t+1}} + \mu_{it}, \tag{3.8}$$

$$\mu_{it} \geq 0 \perp h_{i,t+1} \geq 0, \tag{3.9}$$

where λ_{it} denotes the Lagrange multiplier on the transition of permit holding (3.2) and μ_{it} denotes the Lagrange multiplier on the non-borrowing constraint (3.3). Note that λ_{it} is interpreted as the shadow price of emissions permits for firm i.

Equation (3.6) determines the optimal production decision given the shadow costs of emissions permits. The left-hand-side is the marginal profit that accounts for the cost associated with emissions $\lambda_{it}\rho_{jt}$. This condition can be also written as

$$\frac{P_t^{elec} - c_{jt}^{fuel} - g'(q_{jt})}{\rho_{jt}} = \lambda_{it},$$

implying that the marginal profit from additional emissions should be equal to the shadow costs of emissions permits λ_{it}.

Equations (3.7) and (3.9) determine the shadow costs λ_{it} and μ_{it} from the trading and banking decisions. Equation (3.7) says that the shadow price is equal to the sum of the market price and the marginal trading costs $TC'(b_{it})$. Equations (3.8) and (3.9) show that the shadow value of an emissions permit today is equal to the sum of the discounted marginal value of holding an additional permit tomorrow and the shadow value of borrowing (when it is binding). These conditions along with the transition equation of permit holdings determine the optimal choices for production $\{q_{jt}\}_j$, trading b_{it}, and the banking $h_{i,t+1}$.

The optimality conditions for the non-trader are the same as above except we do not have equation (3.7), and $b_{it} = 0$. These conditions implies that the shadow value of an emissions permit is not directly related to today's permit price in this case. Rather, the shadow value is given by the discounted marginal value from equation (3.8).

Next, I consider the participation decision. If a firm has no prior trading experience (i.e., $I_{it} = 0$), it can choose whether to participate in the market by paying $F_{it}(= F + \epsilon_{it})$. The optimal participation decision is given by

$$D_{it} = \mathbf{1}\left\{V_{it}^1(h_{it}, R_{it}) - (F + \epsilon_{it}) > V_{it}^0(h_{it}, R_{it})\right\},$$

and the participation probability is

$$\mathbb{P}_{it}(h_{it}, R_{it}) = \int \mathbf{1}\left\{V_{it}^1(h_{it}, R_{it}) - (F + \epsilon_{it}) > V_{it}^0(h_{it}, R_{it})\right\} dG(\epsilon_{it}).$$

If a firm already participated in trading (i.e., $I_{it} = 1$), it does not have to pay the participation costs.

Based on the optimal choices for traders and non-traders, I now provide the value function. Let $V_{it}(h_{it}, I_{it}, R_{it}, \epsilon_{it})$ be the value function after observing the random draw of the participation costs. The value function is given by

$$V_{it}(h_{it}, I_{it}, R_{it}, \epsilon_{it}) = \begin{cases} \max\left\{V_{it}^0(h_{it}, R_{it}), V_{it}^1(h_{it}, R_{it}) - (F + \epsilon_{it})\right\} & if\ I_{it} = 0 \\ V_{it}^1(h_{it}, R_{it}) & if\ I_{it} = 1. \end{cases}$$

Also, the ex-ante value functions (before observing ϵ_{it}) are

$$EV_{it}(h_{it}, I_{it}, R_{it}) = \begin{cases} \int \max\left\{V_{it}^0(h_{it}, R_{it}), V_{it}^1(h_{it}, R_{it}) - (F + \epsilon)\right\} dG(\epsilon) & if\ I_t = 0 \\ V_{it}^1(h_{it}, R_{it}) & if\ I_t = 1. \end{cases}$$

By applying the Williams-Daly-Zachary theorem and the envelope theorem (see Appendix C.1 for the derivation), the derivative of the expected value function with respect to the state variable h_{it} can be expressed as follows:

$$\frac{dEV_t(h_{it}, 0, R_{it})}{dh_{it}} = \mathbb{P}_{it}(h_{it}, R_{it})\lambda_{it}^1 + (1 - \mathbb{P}_t(h_{it}, R_{it}))\,\lambda_{it}^0. \tag{3.10}$$

$$\frac{dEV_t(h_{it}, 1, R_{it})}{dh_{it}} = \lambda_{it}^1, \tag{3.11}$$

where λ_{it}^1 and λ_{it}^0 are the Lagrange multipliers on the transition constraint in the optimization problems (3.4) and (3.5), respectively.

Continuation Value at the Terminal Period My model has a finite time period, and the terminal period T corresponds to the year 2003, which is the last period of my sample. However, the cap-and-trade program continued after 2003, and the banking at the end of 2003 was still substantial in the data. To deal with this issue, I introduce the reduced-form continuation value function $CV_{T+1}(h_{i,T+1}, R_i^2)$ in the model. This term captures the banking incentive at the terminal period T. I parametrize the functional form of $CV_{T+1}(h_{i,T+1}, R_i^2)$ and and estimate it along with the other parameters.

3.5 Investment Decisions on Emissions Rate

I now introduce the investment decision on abatement options. In my model, a firm determines phase-specific emissions rates at the beginning of each phase:

$$R_{it} = \begin{cases} R_i^1 & t = 1995, \cdots, 1999 \\ R_i^2 & t = 2000, \cdots, 2003 \end{cases}.$$

The lower the emissions rate, the higher the level of investment in my model. I also assume that the emissions rate is a continuous choice variable. I denote the cost function of investment by $\Gamma(\bar{R}-R)$, where R is the emissions-rate level a firm chooses and $\bar{R}$ is the emissions rate before the investment.

The investment problem for Phase I is given by

$$\begin{aligned} \max_{R_{i,P1}} \quad & EV_{i,1995}(0,0,R_i^1) - \Gamma(R_i^0 - R_i^1) \\ \text{s.t.} \quad & R_i^1 \leq R_i^0, \end{aligned} \tag{3.12}$$

where R_i^0 is the emissions rate in 1990, that is, before the regulation. I incorporate the adjustment costs and the irreversibility of investment by allowing R_i^0 to affect both the investment cost and the choice set of emissions rate R_i^1. Note that $h_{i,1995} = 0, I_{i,1995} = 0$ by definition.

The problem for Phase II is similarly defined as

$$\begin{aligned} \max_{R_{i,P2}} \quad & EV_{i,2000}(h_{i,2000}, I_{i,2000}, R_i^2) - \Gamma(R_i^1 - R_i^2). \\ \text{s.t.} \quad & R_i^2 \leq R_i^1 \end{aligned} \tag{3.13}$$

The investment cost now depends on R_i^1, which is endogenously determined in Phase I.[15]

3.6 Dynamic Competitive Equilibrium with Perfect Foresight

I now define an equilibrium for the permit market. I assume that firms have perfect foresight over the future environment and the only stochastic shock is the participation cost ϵ_{it}. [16]

Definition 1. In a finite-period competitive equilibrium with perfect foresight, a sequence of permit prices $\{P_t\}_{t=1995}^{2003}$ is determined such that

(1) [Optimization] Each firm i optimally chooses $\{\{q_{jt}^*\}_j, b_{it}^*, h_{i,t+1}^*\}_{t=1995}^{2003}$ and $\{R_i^{1*}, R_i^{2*}\}$ given a sequence of permit prices, and

(2) [Market Clearing] $\sum_i b_{it}^* + \bar{B}_t^{fringe}(P_t) = 0$ for $t = 1995, \cdots, 2003$.

I plan to provide a formal argument for the existence of a dynamic competitive equilibrium. My

[15] Four types of firms exist: (1) those that own coal units in only the Phase I group, (2) those that own coal units in Phase I and II groups, (3) those that own coal units in only the Phase II group, and (4) those that only own gas or oil units. Equations (3.12) and (3.13) are the problems for the first type of firms. For the second type of firms, the pre-investment emissions level in Phase II is given by $\omega R_i^1 + (1-\omega)R_i^{0,Phase2}$, where $R_i^{0,Phase2}$ is the emissions rate of Phase II units in 1990 and ω is the share of Phase I units in firm i in terms of generation capacity. In the case of the third type of firm, the pre-investment emissions level in Phase II is $R_i^{0,Phase2}$. The last type of firms does not choose the emissions rate in the model.

[16] Incorporating aggregate uncertainty (i.e., allowing aggregate state variables such as fuel price to stochastically evolve) is a challenging task. Under this setting, the permit price P_t also becomes a random variable, and firms have to form an expectation over future price. Specifying a form of expectation is a major challenge because the permit price itself is an equilibrium object in my model. This setting is close to Krusell and Smith (1998) on a heterogeneous macro model, Cullen (2015) on a dynamic competitive equilibrium in electricity competition, Lee and Wolpin (2006) on structural estimation of a general equilibrium labor model, and Richards-Shubik (2015) on structural estimation of a peer-effect model.

model is close to models of a dynamic competitive market as in Jovanovic (1982), Hopenhayn (1990), Hopenhayn (1992), and Cullen and Reynolds (2017). They show the existence of an equilibrium by providing a correspondence between the social planner's solution and a competitive equilibrium. Regarding uniqueness of equilibrium, I try different initial prices of emissions permits when I numerically solve a dynamic competitive equilibrium and find that these initial values converge to the same equilibrium prices. I leave a formal proof of equilibrium uniqueness to future revisions.

3.7 Model Implications

I discuss several implications from my structural model.

3.7.1 Role of Transaction Costs

I discuss three implications of transaction costs $TC(\cdot)$ and F that I introduced in the model: (i) the efficiency property of a cap-and-trade program, (ii) the independence property (Coase theorem), and (iii) dynamic implications.

To discuss these implications, I first provide optimality conditions whereby no transaction costs exist, namely, $T(\cdot) = 0$ and $F_{it} + \epsilon_{it} = 0$. In this case, equation (3.7) imply that $\lambda_{it} = P_t$ holds for all i, meaning all the firms have the same shadow price, which is given by the market price. Also, equation (3.8), along with envelope conditions (3.10) and (3.11), implies that $\frac{dEV_{i,t+1}(h_{i,t+1}, I_{i,t+1})}{dh_{i,t+1}} = P_{t+1}$. Summarizing these optimality conditions , we have

$$\begin{aligned} \frac{\tau_t^{elec} - c_{jt}^{fuel} - \frac{\partial g(q_{jt},k_j)}{\partial q_{jt}}}{\rho_{jt}} &= P_t. && (3.14) \\ P_t &= \beta P_{t+1} + \mu_{it} && (3.15) \\ & \mu_{it} \geq 0 \perp h_{i,t+1} \geq 0. \end{aligned}$$

I now explain the three implications in turn.

Efficiency Property One of the virtues of cap-and-trade regulation is that the allocation of emissions, given the emissions cap, is efficient in the absence of transaction costs. This assertion is confirmed by equation (3.14)'s implication that the marginal profit from producing one unit of emissions is equalized across firms at the level of permit price P_t. The key mechanism is that all firms are facing the same shadow value given by the market price P_t.

I now examine how the trading behavior affects the shadow costs of emissions permits and thus creates an inefficient allocation of emissions in the presence of transaction costs. Consider the case in which three firms exist: one is a buyer (i.e., $b_{buyer,t} > 0$), the other is a seller (i.e., $b_{seller,t} < 0$), and the last one is a non-trader. Equation (3.6) implies that

$$\begin{aligned}
\lambda_{buyer,t} &= P_t + TC'(b_{buyer,t}) > P_t \\
\lambda_{seller,t} &= P_t + TC'(b_{seller,t}) < P_t, \\
\lambda_{nontrader,t} &= \beta \frac{\partial EV_{t+1}(h_{nontrader,t+1}, I_{nontrader,t+1}=0)}{\partial h_{nontrader,t+1}} + \mu_{nontrader,t}.
\end{aligned}$$

The inequalities in the first two lines hold because $TC'(b) > 0$ for $b > 0$ and $TC'(b) < 0$ for $b < 0$. Intuitively, in the presence of variable transaction cost, buyers have to pay additional costs to purchase emissions permits. By contrast, the revenue from selling a unit of emissions permits is the market price minus the marginal transaction costs. Thus, the marginal profit of emissions for the buyer is strictly higher than that for the seller. In other words, buyers produce less and sellers produce more than the efficient level at which the marginal profits of two firms are equalized. When a firm does not trade, the shadow cost of emissions permits is given by the discounted marginal value of permits tomorrow.

Independence Property Another important property of a cap-and-trade program without any frictions (e.g., transaction costs) is that how the regulator allocates emissions permits has no effect on the pattern of emissions in an equilibrium. In equation (3.14), the initial allocation of permits a_{it} and permit holding h_{it} has no role in determining the production q_{jt}. This property is called the independence property of initial allocation of emission permits, which is also known as the Coase theorem.

Once I introduce transaction costs, the independence property no longer holds: regardless of whether a firm participates in trading, h_{it} increases e_{it} and h_{it+1} and decreases b_{it} when a firm participates in trade. [17] Consider first the case in which a firm does not trade. When the permit holding h_{it} increases marginally, the volume of banking $h_{i,t+1}$ also increases, which lowers the discounted marginal value of permit holding in the next period. Because a firm now has a lower shadow cost of emissions, it has an incentive to produce more emissions. Next, consider the case in which a firm participates in the permit market. An increase in permit holding decreases the volume of net purchase b_{it}. Because the variable transaction cost $TC(b_{it})$ is convex in b_{it}, the decrease in b_{it} lowers the marginal transaction costs a firm faces at the margin. Thus, the emission level e_{it} increases. [18] I discuss the detailed derivation in Appendix C.2.

Dynamic Implications Dynamic implications are also different once I introduce transaction costs. I first explain the case without any transaction costs as a benchmark case. Equation (3.15) implies that the equilibrium permit prices P_t should increase at the rate of β^{-1} over time as long as banking volume is positive and no transaction costs exist. This property is known as the Hotelling $r-$percent rule, in which the price of exhaustible resources should increase at the rate of the inverse

[17] Note that both a_{it} and h_{it} have the same implication on the endogenous variables $\{e_{it}, b_{it}, h_{i,t+1}\}$ in period t.

[18] Stavins (1995) discusses the effect of a_t on emissions level e_t in other functional forms of the transaction cost function in a static model of emissions trading.

of the interest rate (see, e.g., Rubin, 1996). Another important implication is that the model does not pin down the individual optimal behavior for trading b_{it} and banking $h_{i,t+1}$ in the absence of the transaction costs, because the discounted marginal value from banking is constant and given by βP_{t+1}, which is equal to the current shadow value P_t in an equilibrium. Thus, marginal values of net purchase b_{it} and banking $h_{i,t+1}$ are always the same, and any choices are equivalent for individual firms as long as a firm can produce the level of emissions given by the optimality condition on production quantity.

I now consider the case when transaction costs are present. Combining optimality conditions 3.7 and 3.8 and using envelope condition 3.11, I obtain the following condition:

$$P_t + TC'(b_{it}) = \beta \left\{ P_{t+1} + TC'(b_{it+1}) \right\} + \mu_{it},$$

which implies that the permit price does not necessarily increase at the rate of β^{-1}.

More importantly, the marginal values of net purchase b_{it} and banking $h_{i,t+1}$ are no longer constant in this setting. The marginal cost from net purchase is increasing due to the convex transaction costs $TC(b_{it})$. Intuitively, buying more permits becomes more difficult. The discounted marginal value from banking is decreasing because it is given by $\beta\{P_{t+1} + TC'(b_{it+1})\}$, and $b_{i,t+1}$ decreases in $h_{i,t+1}$. In other words, the value of permit holding in the next period is not constant, because firms have to pay the transaction costs so that their marginal revenue from selling is decreasing as they try to sell more. These observations imply that the model now pins down the optimal decisions on both net purchase $b_{it}$and banking volume $h_{i,t+1}$.

3.7.2 Incentives to Invest

I now examine how the incentive to invest is determined in my model. Using the envelope theorem, I calculate the marginal returns from reducing emissions rate R_{P1} as follows:

$$-\frac{\partial EV_{1995}}{\partial R^1} = \sum_{t=1995}^{1999} \beta^{t-1995} \left(\lambda_{it} \cdot \sum_j HR_{jt} q^*_{jt} \right) + \sum_{t=1995}^{1999} \beta^{t-1995} \left(\sum_j \frac{\partial c_{jt}}{\partial R^1} q^*_{jt} \right)$$
$$+ \beta^{2000-1995} \frac{\partial}{\partial R^1} \Gamma(R^1 - R^2).$$

The first component is the returns from reducing emissions evaluated at the shadow value λ_{it}. The second component is the additional costs of using a cleaner fuel. Note that $\frac{\partial c_{jt}}{\partial R^1} < 0$ because fuel costs are higher for low-sulfur coals. The last component is the saving of investment costs in Phase II by investment in Phase I.

The primary component in the returns from investment is the first term. By reducing the emissions rate, the firm can marginally reduce emissions by $\sum_j HR_{jt} q^*_{jt}$. This marginal abatement is evaluated at the shadow value of λ_{it}. The returns from investment is thus given by the discounted sum of the returns from marginal abatement. The path of shadow values λ_{it} is key for investment incentives. As I discussed above, λ_{it} and equilibrium permit price P_t are affected by both banking and transaction costs.

First, the path of the shadow prices will be smoothed over the periods when banking is allowed, due to the optimality conditions on banking (3.8) that the current shadow value of emissions is the discounted shadow value in the future period. The smoother path of $\{\lambda_{it}\}$ implies that firms would invest more in the periods with generous emissions caps (Phase I) and then invest less later (Phase II) when banking is available.

Second, as I have shown above, the shadow value of emissions permits for buyers is higher than for sellers:

$$\lambda_{buyer} > P_t > \lambda_{seller}.$$

Thus, buyers have a higher incentive to invest, whereas sellers have a lower incentive. Intuitively, under the presence of transaction costs, buyers face higher shadow costs of emissions permits, making them prefer investing in technology rather than buying permits. Sellers, on the other hand, obtain lower revenue due to the transaction costs, which gives them lower incentives to invest.

4 Estimation Strategy

This section introduces the estimation strategy for the structural model. I conduct estimation in three steps. First, I estimate a reduced-form model for the capacity factor based on the optimality condition for production quantity q_{jt}. Using the estimated reduced-form model, I next estimate the variable transaction costs, $TC(b)$, the distribution of the fixed transaction costs, F_{it}, the continuation value at the terminal period, $CV_{T+1}(h_{i,T+1}, R_i^2)$, and costs of abatement investment, $\Gamma(\bar{R}-R)$. Note that I fix the annual discount factor at $\beta = 0.95$ throughout the paper. Finally, I estimate the fringe demand.

4.1 Step 1: Reduced-Form Model for the Capacity Factor

Recall that the FOC for unit-level production quantity q_{jt} is

$$\frac{\tau_t^{elec} - c_{jt}^{fuel} - \frac{\partial g(q_{jt},k_j)}{\partial q_{jt}}}{\rho_{jt}} \quad = \quad \lambda_{it},$$

which can be written as

$$\frac{\partial g(q_{jt}, k_j)}{\partial q_{jt}} = \tau_t^{elec} - c_{jt}^{fuel} - \lambda_{it}\rho_{jt}.$$

I consider the following reduced-form model for the optimal choice of q_{jt}

$$q_{jt} = \frac{\exp(\gamma(\tau_t^{elec} - c_{jt}^{fuel} - \lambda_{it}\rho_{jt}))}{1 + \exp(\gamma(\tau_t^{elec} - c_{jt}^{fuel} - \lambda_{it}\rho_{jt}))} \cdot k_j.$$

The first component of the right-hand side is the capacity factor as a function of the markup of electricity production, $\tau_t^{elec} - c_{jt}^{fuel} - \lambda_{it}\rho_{jt}$. [19]

By transforming the above model, we have

$$\log \frac{cf_{jt}}{1 - cf_{jt}} = \gamma \left(\tau_t^{elec} - c_{jt}^{fuel} - \lambda_{it}\rho_{jt} \right),$$

where $cf_{jt} \equiv q_{jt}/k_j$ is the capacity factor.

In empirical implementation, I use month-level observations instead of year-level observations. Also, the sample includes generation units that are not affected by the SO_2 regulation, because I include the data before 1995 (before the ARP started) as well as data for units that were not affected at that point (e.g., data of Phase II units before 2000). To accommodate these observations, I consider the following form of the regression equation indexed by month m:

$$\log \frac{cf_{jm}}{1 - cf_{jm}} = \gamma \left(\tau_m^{elec} - c_{jm}^{fuel} - \mathbf{1}\{SO_2reg\}_{jt} \cdot \lambda_{it}\rho_{jt} \right) + u_j + u_m + u_{jm}, \tag{4.1}$$

where u_j is a unit fixed effect, u_m is time fixed effects, and u_{jm} is an error term. The dummy variable $\mathbf{1}\{SO_2reg\}_{jt}$ takes the value of 1 if unit j is under the ARP in year t.

In equation (4.1), I observe output price τ_{jm}^{elec}, fuel costs c_{jt}, and emissions rate per production ρ_{jt} in the right-hand-side . However, I cannot directly observe the firm-level shadow costs λ_{it}, which are endogenously determined in the structural model. I thus proxy λ_{it} by using the optimality condition from the model. Equation (3.7) implies that if a firm i already participated in permit trading (i.e., $I_{it} = 1$ or $D_{it} = 1$), λ_{it} is given by

$$\lambda_{it} = P_t + TC'(b_{it}).$$

Here I consider a quadratic specification of $TC(b_{it})$ and specify $TC'(b_{it})$ as a linear function: $\theta_0 b_{it} + \theta_i size_i b_{it}$, where $size_i$ is the firm size measured by the sum of the generation capacity of firm i. Then, λ_{it} is written as $\lambda_{it} = P_t + \theta_0 b_{it} + \theta_1 size_i b_{it}$. Putting this equation into equation (4.1), I obtain

$$\begin{aligned} \log \frac{cf_{jm}}{1 - cf_{jm}} \quad = \quad & \tilde{\theta}_1(\tau_m^{elec} - c_{jm}^{fuel}) + \tilde{\theta}_2 \mathbf{1}\{SO_2reg\}_{jt} P_t \rho_{jt} \\ & + \tilde{\theta}_3 \mathbf{1}\{SO_2reg\}_{jt} b_{it}\rho_{jt} + \tilde{\theta}_4 \mathbf{1}\{SO_2reg\}_{jt} \cdot size_i b_{it}\rho_{jt} \\ & + u_j + u_m + u_{jm}, \end{aligned}$$

where $\tilde{\theta}_1 = \gamma, \tilde{\theta}_2 = -\gamma, \tilde{\theta}_3 = -\gamma\theta_0$, and $\tilde{\theta}_4 = -\gamma\theta_1$.

The remaining concern is the endogeneity of ρ_{jt} and b_{it}, both of which are choice variables in the structural model. I use the pre-regulation SO_2 emissions rate $\rho_j^{1990} \equiv HR_j \cdot R_{j,1990}$ as an instrument for ρ_{jt}, where $R_{j,1990}$ is the emissions rate in 1990. Another instrument is the sum of

[19]The reduced-form model can be derived from the following functional form: $g(q,k) = \frac{1}{\gamma}\left(q\log(q) + (k-q)\log(k-q)\right)$.

initial permit allocation for other units within the same firm, $\sum_{k \in J_i, k \neq j} a_{kt}$. I exclude the initial allocation of unit j because the unit-level allocation might depend on the unobserved characteristics of that unit. I use two-stage least squares to estimate the above parameters.

4.2 Step 2: Estimation of Remaining Parameters

Estimation in step 1 gives me the profit function $\pi_{it}\left(\{q_{jt}\}_j\right)$. The next step is to estimate the remaining parameters including transaction costs, the continuation value, and investment costs. I first provide specifications for these primitives.

My model contains two type of transaction costs: variable costs and participation costs. The variable transaction cost function $TC(|b|)$ is specified as follows:

$$TC(|b|) = \frac{\exp(\eta_0 + \eta_1 \log(size_i))}{2} b^2,$$

where $size_i$ denotes firm i's size measured by the sum of the generation capacity of firm i. Participation cost F_{it} is specified as $F_{it} = F + \epsilon_{it}$ where ϵ_{it} follows i.i.d. type 1 extreme value distribution with standard deviation σ_F.

I consider the following parameterization of the continuation value in the terminal period:

$$CV(h_{i,T+1}, R_i^2) = \exp\left(\alpha_0 + \alpha_1 \log(size_i) + \alpha_2 R_i^2\right) h_{i,T+1}^{\alpha_3}.$$

The coefficient depends on the firm size, $size_i$, and the emissions rate in Phase II, R_i^2. These variables capture the heterogeneity in the incentives to bank in the terminal period.

The specification for the investment cost $\Gamma(\cdot)$ is given by

$$\Gamma(\bar{R} - R) = \frac{exp(\zeta_0 + \zeta_1 \log(K_{i,\tau}))}{2} (\bar{R} - R)^2,$$

where $K_{i,\tau}$ is the generation capacity of units regulated in Phase $\tau \in \{I, II\}$. The parameters I estimate in this step are summarized by $\boldsymbol{\theta} = (\eta_0, \eta_1, F, \sigma_F, \alpha_0, \alpha_1, \alpha_2, \alpha_3, \zeta_0, \zeta_1)$.

I use simulated nonlinear least squares to estimate the model parameters. For a given candidate of parameter $\boldsymbol{\theta}$, I solve the model to obtain the prediction of choice variables and match the prediction with the data. However, solving a dynamic competitive equilibrium for each candidate of parameters is computationally infeasible. Instead, I use the observed prices of emissions permits as an equilibrium price and solve only the individual optimization problems to get the model predictions of individual behavior.

This estimation approach builds on the literature of estimation of dynamic structural models in industrial organization and labor economics.[20] My structural model of a cap-and-trade regulation belongs to the class of dynamic competitive equilibrium models. Unlike Lee and Wolpin (2006),

[20] Examples are Rust (1987), Hotz and Miller (1993), and Aguirregabiria and Mira (2002) for a single-agent dynamic discrete choice model, Lee and Wolpin (2006) for a dynamic competitive equilibrium model, and Aguirregabiria and Mira (2007) and Bajari et al. (2007) for dynamic Markov games. See Aguirregabiria and Mira (2010) for a survey of this literature.

however, I can avoid solving for a dynamic competitive equilibrium in estimation, because I can use the observed prices of emissions permits as a sequence of equilibrium prices. The observed prices are fed into solving the individual optimization problems, which are much easier to solve than the dynamic competitive equilibrium, to construct the objective function in estimation. [21]

The procedure of obtaining the model prediction is as follows.

1. Fix a candidate of parameter $\boldsymbol{\theta}$ and the observed permit prices $\{P_t\}_{t=1995}^{2003}$.

2. For each firm i , solve the optimization problem by backward induction and obtain the policy functions for emissions $e_{it}(h_{it}, I_{it}, R_{it})$, trading $b_{it}(h_{it}, I_{it}, R_{it})$, and banking $h_{it+1}(h_{it}, I_{it}, R_{it})$, the participation probability $P_{it}(h_{it}, R_{it})$, and the investment decisions $R_i^1(h_{i1995}, I_{i1995})$, $R_i^2(h_{i,2000}, I_{i,2000}, R_i^1)$.

3. Using the policy functions, simulate the optimal decisions for each pattern of participation in permit trading. I denote the year of participation by $s \in \{\emptyset, 1995, \cdots, 2003\}$, where $s = \emptyset$ means that a firm does not trade at all in the period. Denote the optimal decision for pattern s by $\hat{x}_{it}(s)$.

4. Calculate the probability that each pattern of participation timing is realized. Denote this probability by $Prob_{it}^{enter}(s)$.

5. The prediction for firm i in year t is then given as

$$\hat{x}_{it} = \sum_{s \in \{\emptyset, 1995, \cdots, 2003\}} Prob_i^{enter}(s) \hat{x}_{it}(s). \tag{4.2}$$

Using the simulated choices, I construct the following objective function given by

$$J(\boldsymbol{\theta}) = J_1(\boldsymbol{\theta}) + J_2(\boldsymbol{\theta}).$$

The first component $J_1(\boldsymbol{\theta})$ is minimizing the distance between the prediction and the data at the firm-and-year level

$$J_1(\boldsymbol{\theta}) = \sum_{i=1}^{N} \left(\mathbf{x}_i^{data} - \hat{\mathbf{x}}_i(\boldsymbol{\theta}) \right)' \Omega_i \left(\mathbf{x}_i^{data} - \hat{\mathbf{x}}_i(\boldsymbol{\theta}) \right),$$

where

$$\mathbf{x}_i^{data} = \left(e_{i,t_{i1}}, \cdots, e_{i,2003}, b_{i,t_{i1}} \cdots, b_{i,2003}, h_{i,t_{i1}+1} \cdots, h_{i,2004}, D_{i,\emptyset}, D_{i,t_{i1}} \cdots, D_{i,2003} \right)$$

[21] This empirical strategy is similar in spirit to that in the two-step estimation of a dynamic game, in which the equilibrium objects are directly recovered from the observed data. For example, Aguirregabiria and Mira (2007) estimates players' beliefs over other players' policies from the observed data and solve the optimal response of a player given the estimated beliefs to construct the pseudo-likelihood function.

and $\hat{\mathbf{x}}_i(\boldsymbol{\theta})$ is the corresponding vector for the model prediction given parameter $\boldsymbol{\theta}$. t_{i1} is the first year of observations for firm i; that is $t_{i1} = 1995$ if firm i owns Group I units (those that are under the ARP from 1995), and $t_{i1} = 2000$ if firm i owns only Group II units (those that are under the ARP from 2000). $D_{i,s}$ is a dummy variable that indicates the timing of firm i's participation; $D_{i,s} = 1$ if firm i enters in year s, and, otherwise, 0; and $D_{i,\emptyset} = 1$ if firm i does not trade at all in the sample period. The weighting matrix Ω_i is a diagonal matrix to adjust for differences in scaling.[22]

The second component $J_2(\boldsymbol{\theta})$ incorporates the market clearing conditions, and is given by

$$J_2(\boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1995}^{2003} \left(\sum_i \hat{b}_{it} - \sum_i b_{it}^{data} \right)^2 .$$

This component of the objective function requires that the estimated parameters are such that the observed prices are close to clearing the market in each period. Note that the sum of net purchases in the data $\sum_i b_{it}^{data}$ may not necessarily be equal to zero, since the sample does not cover all the firms participating in the permit trading.

Standard errors are calculated by bootstrap at the firm-history level. I randomly draw samples of 85 firms with replacement and construct 40 bootstrap samples.

4.3 Step 3: Estimation of Fringe Demand

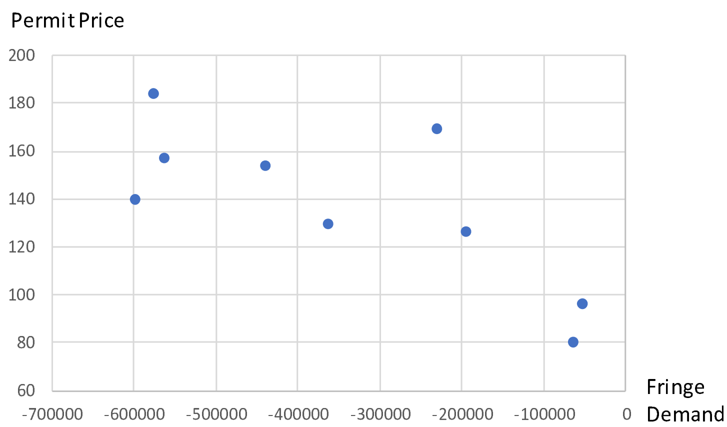
I now estimate the fringe-demand function. I consider the following specification with constant elasticity:

$$\begin{aligned} \bar{B}_t^{fringe} &= -\alpha_t P_t^{\epsilon} \\ \Longleftrightarrow \log\left(-\bar{B}_t^{fringe}\right) &= \epsilon \log P_t + \log(\alpha_t). \end{aligned}$$

One of the difficulties in estimating the fringe function is the paucity of data points. I have only 9 data points, because the sample is between 1995 to 2003, as shown in Figure 7, where I plot permit prices against the fringe demand for each year. I estimate the model by OLS and IV. I use the sum of initial allocations for firms in my sample as an instrument for P_t. I plan to provide a robustness check by using different levels of price elasticity and using different specifications for the fringe function (including a linear specification and a semi-log specification).

[22] The weighting matrix is set as $\Omega_i \equiv \mathrm{diag}(\underbrace{\hat{\mathbf{V}}(e), \cdots, \hat{\mathbf{V}}(e)}_{T_i}, \underbrace{\hat{\mathbf{V}}(b), \cdots, \hat{\mathbf{V}}(b)}_{T_i}, \underbrace{\hat{\mathbf{V}}(h), \cdots, \hat{\mathbf{V}}(h)}_{T_i}, \underbrace{\hat{\mathbf{V}}(D), \cdots, \hat{\mathbf{V}}(D)}_{T_i+1}))$, where $\hat{\mathbf{V}}(e)$, $\hat{\mathbf{V}}(b)$,$\hat{\mathbf{V}}(h)$, and $\hat{\mathbf{V}}(D)$ are sample variances of emissions, net purchase, banking, and participation in the sample, respectively.

Figure 7: Plot of Permit Prices P_t and Fringe Demand $\bar{B}_t^{fringe}$



5 Estimation Results

5.1 Parameter Estimates

Table 2 presents the parameter estimates of the structural model. The model incorporates two types of transaction costs: participation and variable costs. Regarding the variable transaction costs, the coefficient on the firm size is negative but small. This estimate implies that although bigger firms tend to have lower transaction costs, heterogeneity across firms is negligible. Based on the parameter estimates, I calculate the marginal transaction cost, given by $exp(\eta_0 + \eta_1 \log(size_i))|b_{it}|$. The mean of the costs is \$193 and the median is \$98. Considering that the range of the permit prices is within \$100 to \$200 , these numbers are substantial. The mean of the fixed costs of participation is around \$38,000, which is quite small. The parameters in the continuation value function at the terminal period have signs that are intuitively reasonable. Bigger and dirtier firms obtain a higher value from banking. Estimates of the investment cost also have the reasonable sign. The bigger the capacity, the higher the investment costs.

Table 2: Parameter Estimates

	Parameter	Description	Estimate	Standard Errors
Production Parameters: $g(q,k) = \frac{1}{\gamma}\left(q\log(q) + (k-q)\log(k-q)\right)$	γ	Curvature	4.333e-03	1.149e-03
Variable Costs: $TC(b) = \frac{1}{2}exp(\eta_0 + \eta_1 log(size))\|b\|^2$	η_0	Constant	-3.516	0.343
	η_1	Firm size	-0.011	0.011
Participation Costs: F_{it} $logit(F, \sigma_F)$	F	mean (1 USD)	38,379	148,846
	σ_F	SD (1 USD)	1,259	12,405,130
Continuation Value:$CV(h_{i,T+1}, R_i^2)$ $\exp(\alpha_0 + \alpha_1\log(size_i) + \alpha_2 R_i^2)h_{i,T+1}^{\alpha_3}$	α_0	Constant	4.165	2.267
	α_1	Firm Size	0.898	0.197
	α_2	Emissions Rate	0.036	0.183
	α_3	Curvature	0.314	0.053
Investment Costs:$\Gamma(R_{i,\tau}, \bar{R}_{i,\tau})$ $\frac{1}{2}exp(\zeta_0 + \zeta_1\log(K_{i,\tau}))(\bar{R}_{i,\tau} - R_{i,\tau})^2$	ζ_0	Constant	11.725	1.137
	ζ_1	Capacity	0.697	0.251

Table 3 shows the estimates of the fringe elasticity. In OLS, the elasticity is estimated to be 2.72, while the estimate in the IV specification is 2.39. In the counterfactual analysis, I use 2.39 as a benchmark parameter. I plan to provide a robustness check on this parameter.

Table 3: Parameter Estimates on Fringe Demand

	[1] log-log		[2] log-log		[3] linear		[4] linear	
Price	2.72	(0.52)	2.39	(0.77)	4386.5	(1184.91)	5106.0	(1275.42)
Phase II dummy	0.34	(0.33)	0.40	(0.43)	83197.5	(108217)	68423.0	(108363)
Constant	-0.99	(2.58)	0.57	(3.54)	-297448.9	(150911)	-389813.9	(170423)
Method	OLS		IV		OLS		IV	

Note: Standard errors are shown in the brackets.

5.2 Model Fit

This subsection discusses the model fit under the estimated parameters. I first solve a dynamic competitive equilibrium under the estimated parameters and obtain model predictions on permit prices and individual behavior. Appendix A explains the algorithm for solving a competitive equilibrium in detail.

Figure 8 shows the predicted and observed prices in the data. The blue dashed line corresponds to the data and the orange real line to the model prediction. Although the path of the predicted equilibrium prices deviates from the observed one in several years, 1996, 1997, and 1999, my model predicts the price quite well in Phase II (2000-2003). One potential reason for the worse model fit in the early periods is the lack of experiences in permit trading. Once firms accumulate sufficient experiences, the market price of emissions permits in Phase II becomes close to the prices predicted by my equilibrium model.

Figure 8: Model Fit of Equilibrium Permit Prices

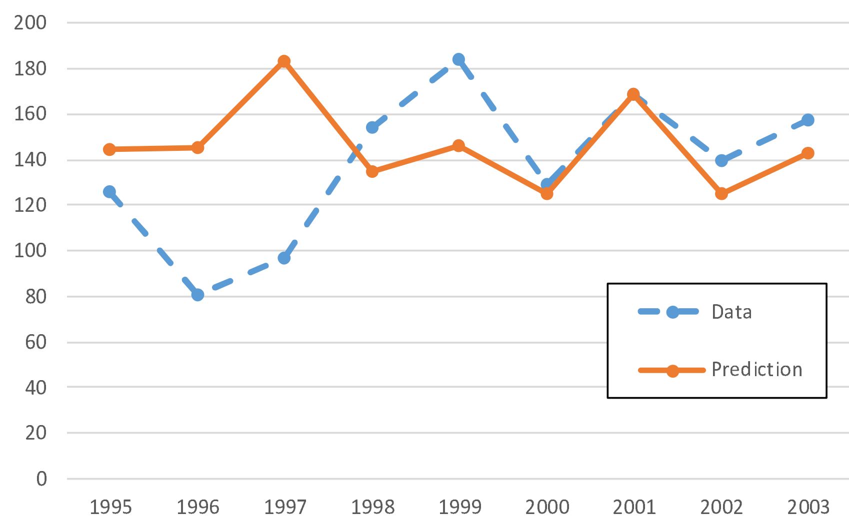


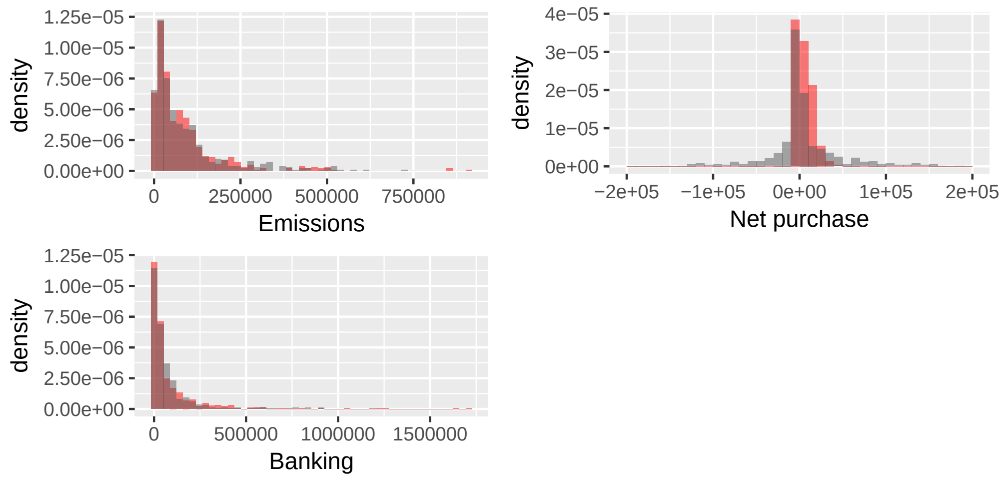
Table 4 and Figure 9 show the model fit in terms of individual behavior under the predicted equilibrium prices. The predictions on emissions e_{it} and banking h_{it+1} are quite close to the observed distribution. Although both lower and upper tails of the distribution of net-purchase are different between the prediction and the data, my model predicts the median and mean net purchase quite well.

In the counterfactual simulations in section 6, I refer to the equilibrium outcome I solved here as a baseline outcome.

Table 4: Model Fit

	Emissions e_{it}		Net Purchase b_{it}		Banking $h_{i,t+1}$	
	Data	Prediction	Data	Prediction	Data	Prediction
Min	849	0	-335,072	-55,375	7	0
1st quantile	18,685	20,232	-3,916	-1,440	7,652	6,751
Median	47,283	48,222	0	2,131	30,283	24,487
Mean	89,596	85,116	5,764	5,411	75,251	83,183
3rd quantile	109,114	104,339	12,652	11,470	79,356	84,081
Max	727,040	910,781	351,702	46,039	1,204,817	1,711,434

Figure 9: Histogram of the Model Prediction and the Data



Note: Gray histogram corresponds to the data, and red histogram corresponds to the model predictions.

6 Counterfactual Experiments

This section discusses a series of counterfactual exercises using the model with the estimated parameters from section 5. In section 6.1, I first examine the effect of the cap-and-trade program and decompose it into permit trading and permit banking. I then quantify potential gains from trade by simulating the market outcome in the absence of transaction costs in section 6.2.

In counterfactual simulations, I fix model primitives at the estimated values. This assumption could be problematic for the continuation value at the terminal period $CV_{T+1}(\cdot)$. The continuation value captures the incentive of permit banking at the terminal period, which could change in the counterfactual situations. An alternative approach would be to fix the volume of permit banking at the terminal period, instead of the continuation value function. I will work on this approach in the future revision as a robustness check.[23]

The supply function from fringe firms $\bar{B}_t^{fringe}(\cdot)$ could potentially change in the counterfactual scenarios as well. In the simulation below, I fix the *level* of the fringe supply, instead of the function itself, in each year at the level under the equilibrium I solved in section 5.2. Under this assumption, I fix the total number of emissions permits available for firms in my sample, including their initial allocation and the fringe supply.

[23] Another approach is to model the terminal period as a stationary and infinite-period dynamic programming problem. This approach allows me to avoid specifying a parametric form of the continuation value. However, this approach is computationally more demanding because I have to solve the value function by a contraction mapping for each firm. I plan to take these approaches in the future revision as a robustness check.

6.1 Experiment 1: Effects of a Cap-and-Trade Program and its Decomposition

I simulate the outcomes in the following two situations. First, I simulate the case in which all firms are required to achieve the uniformly determined level of emissions rate in each phase. The target emissions rates are defined as the rates under which the aggregate emissions are equal to the level of the baseline outcome, which I calculated in section 5.2. This simulation outcome corresponds to the one under uniform standard regulation. The difference between this outcome and the baseline outcome quantifies the effects of a cap-and-trade program. Second, I simulate a market outcome whereby permit banking is completely banned. This simulation allows me to decompose the effects of a cap-and-trade program into the effects of permit trading and the effects of permit banking. Appendix B explains how to simulate the equilibrium outcome in each case.

Effects of a Cap-and-Trade in Comparison to Uniform Standards Table 5 presents the results of counterfactual experiments. The numbers shown in the table are the totals in my sample period. The upper panel shows the aggregate levels of emissions, left-over permits, and banking at the terminal period. In the absence of a permit banking system, emissions permits that firms have at the end of the period would expire. I call these permit left-over permits. The lower panel shows the welfare measures including abatement costs and health and environmental damages. I explain the definition of health and environmental damages below.

I first discuss the effects of the cap-and-trade program in comparison to the uniform standard. The simulation results show that the cap-and-trade decreases the total costs of abatement, including investment costs and loss of profit from electricity generation, by 3.1 billion USD in total, or by 16.6%. The cost saving is allowed by a more flexible pattern of investment as shown in Figure 10. In the case of uniform standard, there is a mass of firms at 2.26 lbs/MMBtu in Phase I and 1.31 lbs/MMbtu in Phase II. These numbers are the emissions rates that achieve the same aggregate emissions as the baseline in Phase I and II, respectively.[24] Introduction of the cap-and-trade program allows firms to optimally decide their emissions rate based on the returns and costs from investment, which depend on the characteristics of firms, permit prices, and transaction costs in the permit trading. Their optimal decisions lead to a more heterogeneous distribution of emissions rates and significant cost saving, compared to the uniform standard.

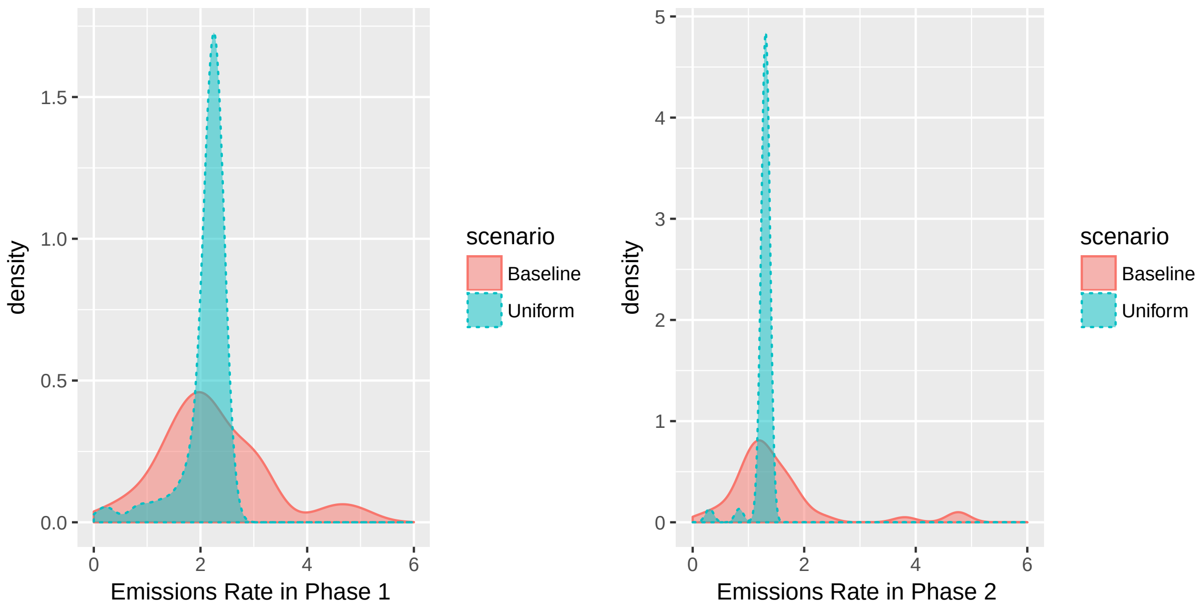
[24]Some firms have the lower emissions rate than the uniform rate in the figure. This is because those firms already had satisfied the uniform standard before the regulation.

Table 5: Effects of a cap-and-trade and its Decomposition

		Baseline	Uniform Standard	No Banking
Emissions (in 1 million ton)		45.54	45.54	41.60
Left-over permits		0.00	n.a.	9.15
Banking at the terminal period		5.21	n.a.	0.00
Firm costs	Investment Costs	15,179	18,648	15,934
(in million USD)	Loss of Electricity Profit	380	0	869
	Total costs	15,559	18,648	16,803
	change from baseline		3,090	1,245
Health and environmental damages	Total	54,787	54,592	50,692
(in million USD)	change from baseline		-195	-4,095
Total Costs	Total costs	70,346	73,241	67,495
(in million USD)	change from baseline		2,895	-2,850
Firm costs / abatement (in USD)		376	450	371
Health and environmental damage / emissions (in USD)		1,203	1,199	1,219

Note: The numbers are the total from 1995 to 2003. The unit of emissions, left-over permits, and banking at the terminal period is 1 million SO_2 tons. The unit of costs are 1 million USD in 2000.

Figure 10: Distribution of Emissions Rate in the baseline and uniform standard



Note: Emissions rate is measured by lbs per MMBtu.

Implications for Health and Environmental Damages A potential concern of a cap-and-trade program is the implication for health and environmental damages. Even though the aggregate level of emissions is fixed, the distribution of emissions under a cap-and-trade might be different from the one under a uniform standard regulation (See, e.g., Muller and Mendelsohn 2009, Fowlie et al. 2012, Fowlie and Muller 2013, and Chan et al. 2015). In particular, SO_2 emissions are non-uniformly mixed pollution; health and environmental damages depend on the location of the emissions' source. To the extent that the damage from a particular location is positively correlated with abatement costs of power plants in that location, a cap-and-trade program would increase damages from emissions in comparison to a uniform standard.

To discuss the net benefit of a cap-and-trade program, I calculate health and environmental damages in each case in Table 5 using the data from Muller and Mendelsohn (2009). They use the AP2 model, an integrated assessment model, to calculate marginal damages from SO_2 emissions at the county level.[25][26] Following Muller and Mendelsohn (2009), I assume that damages are linear in SO_2 emissions. The emissions damage from a particular county is given by the product of the marginal damage and the total SO_2 emissions from electricity plants that locate in the county.[27]

[25] I use the marginal damages from point sources with effective height less than 250 meters.

[26] The AP2 model is used in various papers, including Fowlie and Muller (2013), Chan et al. (2015), and Holland et al. (2016), for calculating health and environmental damages of air pollutants.

[27] One important limitation of the marginal damage data from Muller and Mendelsohn (2009) is that information about which location is hurt by emissions from a particular county is not available. To obtain this information, I need to run the AP2 model to simulate the flow of SO_2 emissions damages from one county to another county. This additional data would allow me to discuss distributional implications of a cap-and-trade program. I leave this analysis

I find that a cap-and-trade program increases damages from SO_2 emissions slightly by 195 million USD (or 0.3%) in comparison to the uniform standard. This percentage increase of 0.3% is within the range of findings in Chan et al. 2015.[28] Overall, the savings of firm costs are large enough to offset the increase in damages, and thus the cap-and-trade program improves the overall welfare.

Effects of Permit Banking The third column in Table 5 shows the outcome in the absence of a permit banking system. The aggregate emissions are different from the baseline due to the presence of left-over permits. Around 18% of permits would expire if a permit banking system is not available. Note that emissions permits might expire in the following two cases. If a firm does not participate in the permit trading, they cannot sell permits and thus the remaining permits must expire. Even though they participate, the marginal revenue from selling permits could be lower than zero due to transaction costs. In such a case, firms do not sell all of the remaining permits.

The total abatement costs are estimated to be 16.8 billion USD in the absence of permit banking. Using this estimate, I decompose the savings of abatement costs into permit trading and permit banking. I find that permit banking decreases the abatement costs by 1.25 billion USD, accounting for around 26% of the effect of cap-and-trade. However, I need to take into account the difference in the aggregate emissions. To do so, I calculate the average abatement costs in each case. I find that the average abatement cost in the absence of transaction cost is 1.3% lower than the baseline case.

To see this, I calculate the firm-and-year level outcomes in the baseline and no-banking case in Table 6. The table shows that the trading volume (i.e., the absolute value of the net purchase, $|b_{it}|$) is higher by 29% in the absence of permit banking than in the baseline, implying that permit trading is more active without the banking. When banking is not allowed, firms have a higher incentive to trade. Because firms cannot rely on their banked permits for compliance, they have to buy permits from other firms. In the case of sellers, they have to discard emissions permits unless they participate in permit trading to sell. Overall, more active trading of emissions permits leads to efficient allocation of emissions permits. The difference in trading patterns is a potential reason why the average abatement costs are not higher than the baseline case, even though permit banking is not allowed.

to future work.

[28] Chan et al. 2015 focus on 2002 to evaluate the net benefit of the Acid Rain Program. They find that the effect of a cap-and-trade program on health and environmental damages is within the range -0.4% to 1.8%, depending on the model specifications and the counterfactual policy in the absence of a cap-and-trade program.

Table 6: Simulation outcomes at the Firm-and-year Level

	Baseline		
	Mean	Std. Dev	Median
Emissions e_{it}	85,116.1	111,237.8	48,221.6
Net purchase b_{it}	5,410.8	9,326.0	2,131.3
Trading volume $\|b_{it}\|$	7,229.0	7,996.9	3,657.0

	No Banking		
	Mean	Std. Dev	Median
Emissions e_{it}	77,753.9	90,146.3	50,683.4
Net purchase b_{it}	5,000.6	12,353.0	0.0
Trading volume $\|b_{it}\|$	9,300.0	9,539.2	5,215.0
Left-over permits	16,698.7	57,228.5	0.0

6.2 Experiment 2: The Potential Gains from Trade

I now examine the implications of transaction costs. As I discussed in section 3.7.1, transaction costs are a source of inefficiency in a cap-and-trade program. Estimates of my structural model suggest that although the transaction costs in the form of participation cost are small, the variable transaction costs are substantial.

In this simulation, I shut down both participation and variable transaction costs (while allowing permit banking) and solve a market equilibrium. This simulation quantifies the potential gains from trade in the absence of transaction costs. This simulation outcome is also interpreted as an outcome when the regulator introduces a centralized exchange for emissions permits and runs an auction to allocate all the emissions permits.

Table 7 shows the counterfactual result along with the baseline result. The upper panel shows that the level of permit banking at the terminal period is lower in the absence of transaction costs than in the baseline. This pattern can be found in other years, as shown in figure 12. Under the presence of transaction costs, firms are less active in permit trading and thus they rather save emissions permits.

Regarding the abatement costs for firms, the table shows that the costs would be lower by 5.8 billion USD in total, or by 37%. Although this partially reflects the higher emissions level in the absence of transaction costs, the average cost of abatement is also lower by 119 USD per SO_2 tons, or by 31.6%. This finding indicates that the "potential" gains from trade, which could be achieved in the absence of transaction costs are significant.

Where does this additional cost saving come from? Figure 11 plots the distributions of emissions rates in the case without transaction costs and the baseline case. The distribution is more dispersed in the absence of transaction costs than in the baseline. Shutting transaction costs down makes firms trade more actively, so that firms are more flexible in their compliance. Some firms that are costly to reduce emissions by themselves are more likely to purchase emissions permits, whereas other firms invest more because their revenues from selling permits become higher once transaction costs are removed.

Another source of cost saving is due to the lower level of permit banking. Figure 12 shows that the level of banking is higher in the baseline than in the absence of transaction costs. When the transaction costs exist, firms prefer to save emissions permits, instead of selling them in the market. Once transaction costs are shut down, firms have a higher incentive to trade permits with other firms, improving the allocative efficiency of emissions permits.

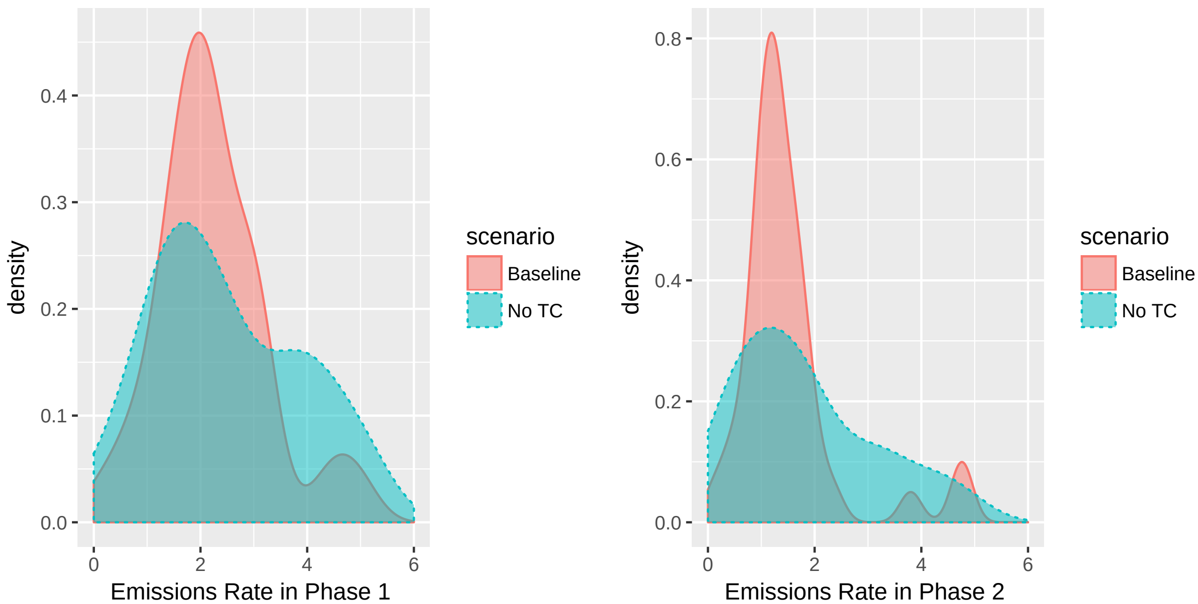
Health and environmental damages increase by 4.5 billion USD in the absence of transaction costs. This increase reflects both the increase in the aggregate level of SO_2 emissions, as well as the increase in the average health damages. In particular, the average health damages increase by 0.7%, indicating that more active trading of emissions permits leads to more emissions in the region where the health damage is higher. In sum, the total costs including firms' costs of abatement and health damages decreased by 1.3 billion USD (1.8%) in the absence of transaction costs.

Table 7: The Potential Gains from Trade

		Baseline	No Transaction Costs
Emissions (in 1 million ton)		45.54	48.99
Dumped permits		0.00	0.00
Banking at the terminal period		5.21	1.76
Firm costs	Investment Costs	15,179	9,740
(in million USD)	Loss of Electricity Profit	380	25
	Total costs	15,559	9,765
	change from baseline		-5,794
Health and environmental damages	Total	54,787	59,321
(in million USD)	change from baseline		4,534
Total Costs	Total costs	70,346	69,086
(in million USD)	change from baseline		-1,259
Firm costs / abatement (in USD)		376	257
Health and environmental damage / emissions (in USD)		1,203	1,211

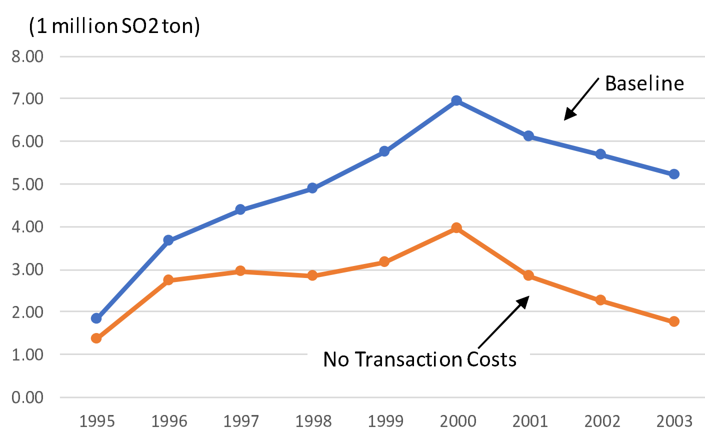
Note: The numbers are the total from 1995 to 2003. The unit of emissions and banking at the terminal period is 1 million SO_2 tons. The unit of costs are 1 million USD in 2000.

Figure 11: Distribution of Emissions Rate in the Absence of Transaction Costs



Note: Emissions rate is measured by lbs per MMBtu.

Figure 12: Permit Banking in the Absence of Transaction Costs



7 Conclusion and Further Directions

I study the welfare effects of a cap-and-trade program on air pollutants in the context of the SO_2 emissions regulation in the US electricity industry. I construct a dynamic equilibrium model of a cap-and-trade program in which firms makes decisions on abatement investment, permit trading, and permit banking. By applying the model to the data from the US Acid Rain Program, I find that variable transaction costs associated with permit trading are substantial.

I use the estimated model to quantify the effects of a cap-and-trade program in comparison to the uniform standard on emissions rate as a counterfactual command-and-control type policy. I find that the cap-and-trade program decreased the aggregate costs of reducing emissions by 340 million USD per year, or 16.6%. Although the damages from SO_2 emissions increased as a result of permit trading, the cost saving is sufficiently high to offset the negative effects on health.

I also simulate the counterfactual outcome in the absence of transaction costs and find that the aggregate costs could be saved further by 643 million USD per year, or by 26%. This additional cost saving is achieved by a more efficient allocation of investment. My findings indicate that the full potential of a cap-and-trade program has not been realized in my sample period.

Several extensions and applications remain for future work. First, it would be interesting to study the implications of policies regarding permit prices in a cap-and-trade program. Recently, the volatile and low permit prices observed in cap-and-trade programs concern policymakers, which leads to the proposal of various measures to stabilize permit prices, including a price floor/ceiling and the Market Stability Reserve. By extending my empirical framework, I could evaluate the effectiveness of these proposed policies and its welfare consequences.

In addition, my framework can be applied beyond a cap-and-trade program on air pollutants. The governments now use a market-based policy in various settings, including credit trading in the CAFE regulation and Renewable Energy Certificates in Renewable Portfolio Standard. Under these policies, firms face a similar problem to the one I study in this paper: a firm can either trade these credits, or invest in technologies (i.e., improving fuel efficiency in the CAFE credit trading, and building renewable generators in the RPS program). My empirical framework can be used to study the effectiveness of these market-based policies and the implications of alternative regulatory designs. I leave these topics for future work.

References

Abito, Jose Miguel, "Agency Costs in Environmental Regulation: Evidence from Regulated Electric Utilities," *Working Paper*, 2014.

Aguirregabiria, Victor and Pedro Mira, "Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models," *Econometrica*, 2002, *70* (4), 1519–1543.

_ **and** _ , "Sequential estimation of dynamic discrete games," *Econometrica*, 2007, *75* (1), 1–53.

_ **and** _ , "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 2010, *156* (1), 38–67.

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 2007, *75* (5), 1331–1370.

Breen, William J, Laurie Simon Hodrick, and Robert A Korajczyk, "Predicting equity liquidity," *Management Science*, 2002, *48* (4), 470–483.

Cantillon, Estelle and Aurelie Slechten, "Price formation in the European carbon market: the role of firm participation and market structure," *Working Paper*, 2015.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L Palmer, "Sulfur dioxide control by electric utilities: What are the gains from trade?," *Journal of political Economy*, 2000, *108* (6), 1292–1326.

Chan, H Ron, B Andrew Chupp, Maureen L Cropper, and Nicholas Z Muller, "The Impact of Trading on the Costs and Benefits of the Acid Rain Program," Technical Report, National Bureau of Economic Research 2015.

Chan, Ron, "How Large are the Cost Savings from Emissions Trading? An Evaluation of the U.S. Acid Rain Program," *Working Paper*, 2015.

Coase, Ronald H, *The problem of social cost*, Springer, 1960.

Collard-Wexler, Allan, "Demand Fluctuations in the Ready-Mix Concrete Industry," *Econometrica*, 2013, *81* (3), 1003–1037.

Cullen, Joseph A, "Dynamic response to environmental regulation in the electricity industry," *Working Paper*, 2015.

_ **and Stanley S Reynolds**, "Market Dynamics and Investment in the Electricity Sector," 2017.

Dardati, Evangelina, "Pollution Permit Systems and Firm Dynamics: How does the Allocation Scheme Matter?," *Working Paper*, 2014.

Dávila, Eduardo and Cecilia Parlatore, "Trading Cost and Informational Efficiency," 2017.

EIA, "Electric Power Annual 2012," Technical Report 2012.

Ellerman, A Denny and Juan-Pablo Montero, "The efficiency and robustness of allowance banking in the US Acid Rain Program," *The Energy Journal*, 2007, pp. 47–71.

_ **, Paul L Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth M Bailey**, *Markets for clean air: The US Acid Rain Program*, Cambridge University Press, 2000.

Ericson, Richard and Ariel Pakes, "Markov-perfect industry dynamics: A framework for empirical work," *The Review of Economic Studies*, 1995, *62* (1), 53–82.

Fabra, Natalia and Mar Reguant, "Pass-Through of Emissions Costs in Electricity Markets," *American Economic Review*, 2014, *104* (9), 2872–99.

Fowlie, Meredith, "Allocating emissions permits in cap-and-trade programs: Theory and evidence," *Working Paper*, 2010.

_ , "Emissions trading, electricity restructing, and investment in pollution abatement," *The American Economic Review*, 2010, pp. 837–869.

_ **and Jeffrey M Perloff**, "Distributing pollution rights in cap-and-trade programs: are outcomes independent of allocation?," *Review of Economics and Statistics*, 2013, *95* (5), 1640–1652.

_ **and Nicholas Muller**, "Market-based emissions regulation when damages vary across sources: What are the gains from differentiation?," Technical Report, Working Paper 2013.

_ **, Mar Reguant, and Stephen P Ryan**, "Market-based emissions regulation and industry dynamics," *forthcoming in Journal of Political Economy*, 2014.

_ **, Stephen P Holland, and Erin T Mansur**, "What do emissions markets deliver and to whom? Evidence from Southern California's NOx trading program," *The American economic review*, 2012, *102* (2), 965–993.

Gârleanu, Nicolae and Lasse Heje Pedersen, "Dynamic trading with predictable returns and transaction costs," *The Journal of Finance*, 2013, *68* (6), 2309–2340.

Helfand, Gloria E, Michael R Moore, and Yimin Liu, "Testing for dynamic efficiency of the Sulfur Dioxide Allowance Market," 2006.

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates, "Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors," *American Economic Review*, 2016, *106* (12), 3700–3729.

Hopenhayn, H. A., "Industry Equilibrium Dynamics: A General Competitive Theory," March 1990.

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Hotz, V Joseph and Robert A Miller, "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 1993, *60* (3), 497–529.

Jaraitė-Kažukauskė, Jūratė and Andrius Kažukauskas, "Do transaction costs influence firm trading behaviour in the european emissions trading system?," *Environmental and Resource Economics*, 2015, *62* (3), 583–613.

Joskow, Paul L and Richard Schmalensee, "The Political Economy of Market-Based Environmental Policy: the US Acid Rain Program 1," *The journal of law and economics*, 1998, *41* (1), 37–84.

_ , _ , and Elizabeth M Bailey, "The market for sulfur dioxide emissions," *American Economic Review*, 1998, pp. 669–685.

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica: Journal of the Econometric Society*, 1982, pp. 649–670.

Kalouptsidi, Myrto, "Time to build and fluctuations in bulk shipping," *The American Economic Review*, 2014, *104* (2), 564–608.

Kellogg, Ryan, "The effect of uncertainty on investment: evidence from Texas oil drilling," *The American Economic Review*, 2014, *104* (6), 1698–1734.

Keohane, Nathaniel, "Cost savings from allowance trading in the 1990 Clean Air Act: Estimates from a choice-based model," *Moving to markets in environmental regulation: Lessons from twenty years of experience*, 2006, pp. 194–229.

Kolstad, Jonathan T and Frank A Wolak, "Using Environmental Emissions Permit Prices to Raise Electricity Prices: Evidence from the California Electricity Market," *Working Paper*, 2008.

Krusell, Per and Anthony A Smith Jr, "Income and wealth heterogeneity in the macroeconomy," *Journal of political Economy*, 1998, *106* (5), 867–896.

Lee, Donghoon and Kenneth I Wolpin, "Intersectoral labor mobility and the growth of the service sector," *Econometrica*, 2006, *74* (1), 1–46.

Lillo, Fabrizio, J Doyne Farmer, and Rosario N Mantegna, "Econophysics: Master curve for price-impact function," *Nature*, 2003, *421* (6919), 129–130.

Liski, Matti and Juan-Pablo Montero, "Market Power in an Exhaustible Resource Market: The Case of Storable Pollution Permits," *The Economic Journal*, 2011, *121* (551), 116–144.

Muller, Nicholas Z and Robert Mendelsohn, "Efficient pollution regulation: getting the prices right," *The American Economic Review*, 2009, *99* (5), 1714–1739.

Reguant, Mar and A Denny Ellerman, "Grandfathering and the endowment effect: An Assessment in the context of the Spanish National Allocation Plan," *Working Paper*, 2008.

Richards-Shubik, Seth, "Peer effects in sexual initiation: Separating demand and supply mechanisms," *Quantitative Economics*, 2015, *6* (3), 663–702.

Robert, Engle, Ferstenberg Robert, and Russell Jeffrey, "Measuring and modeling execution cost and risk," *The Journal of Portfolio Management*, 2012, *38* (2), 14–28.

Rubin, Jonathan D, "A model of intertemporal emission trading, banking, and borrowing," *Journal of Environmental Economics and Management*, 1996, *31* (3), 269–286.

Rust, John, "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, 1987, pp. 999–1033.

_ , "Structural estimation of Markov decision processes," *Handbook of econometrics*, 1994, *4*, 3081–3143.

Ryan, Stephen P, "The costs of environmental regulation in a concentrated industry," *Econometrica*, 2012, *80* (3), 1019–1061.

Schennach, Susanne M, "The economics of pollution permit banking in the context of Title IV of the 1990 Clean Air Act Amendments," *Journal of Environmental Economics and Management*, 2000, *40* (3), 189–210.

Schmalensee, Richard and Robert N Stavins, "The SO2 Allowance Trading System: The Ironic History of a Grand Policy Experiment," *The Journal of Economic Perspectives*, 2013, *27* (1), 103–121.

Stavins, Robert N, "Transaction costs and tradeable permits," *Journal of environmental economics and management*, 1995, *29* (2), 133–148.

U.S. Environmental Protection Agency, "The National Allowance Data Base Version 2.1 Technical Support Document," 1993.

_ , "Technical Documentation for Phase II Allowance Allocations," Technical Report, EPA-430-R-93-002 1993.

A Computational Details on Solving the Structural Model

Appendix A explains the details on the computational procedure of solving the structural model.

A.1 Individual Optimization

I first explain the computational procedure for solving an individual problem. For notational simplicity, I omit the script i for a particular firm. Because the model has a finite period, it can be solved by backward induction.

1. Phase II (2003 to 2000): I solve the optimization problem from 2003 to 2000. Note that I use $CV_{T+1}(h_{T+1}, R^2)$ as a continuation value in the terminal period 2003. By solving in a backward way, I obtain the policy function $\hat{x}_t(h_t, I_t, R^2)$ for emissions e_t, net purchase b_t, and banking h_{t+1}, and the expected value function in 2000 $EV_{2000}(h_{2000}, I_{2000}, R^2)$.

2. Investment decision for Phase II: I define the continuation value at the timing of making the investment decision for Phase II by $W_{2000}(h_{2000}, I_{2000}, R^2)$. The decision problem is given by

$$\begin{aligned} W_{2000}(h_{2000}, I_{2000}, R^1) \equiv \quad & \max_{R^2} \quad EV_{2000}(h_{2000}, I_{2000}, R^2) - \Gamma(R^2, R^1). \\ & \text{s.t.} \quad\;\; R^2 \leq R^1 \end{aligned}$$

   By solving this problem, I obtain the investment policy function $R^{2*}(h_{2000}, I_{2000}, R^1)$.

3. Phase I (1999 to 1995): I repeat the same procedure as step 1. Note that the continuation value in the problem at $t = 1999$ is given by $W_{2000}(h_{2000}, I_{2000}, R^1)$.

4. Investment for Phase I: The problem is given by

$$\begin{aligned} & \max_{R^1} EV_{1995}(0, 0, R_{P1}) - \Gamma(R^1, R^0). \\ & \text{s.t.} R^1 \leq R^0 \end{aligned}$$

   Note that $h_{1995} = 0$ and $I_{1995} = 0$ in 1995.

A.2 Computation of a Dynamic Competitive Equilibrium

The computational procedure for finding an equilibrium is parallel to the estimation procedure that I introduced in section 5.

1. Fix a candidate of permit prices $\mathbf{P} = \{P_t\}_{t=1995}^{2003}$.

2. Solve the individual problem by backward induction and obtain the policy function $\hat{x}_{it}(h_{it}, I_{it}, R_{it})$ for emissions e_t, net purchase b_t, and banking h_{t+1}, participation probability $P_{it}(h_{it}, R_{it})$, and the investment decisions $R_i^1(h_{i1995}, I_{i1995})$ and , $R_i^2(h_{i,2000}, I_{i,2000}, R_i^1)$.

3. Consider the timing of market participation. Denote the year of participation by $s \in \{\emptyset, 1995, \cdots, 2003\}$. $s = \emptyset$ means that a firm does not trade at all in the period.

4. For each path of participation timing, I simulate the optimal decisions using the policy functions.

5. Calculate the probability that each path of participation timing is realized.

6. The simulated optimal decisions are given as

$$\hat{x}_{it} = \sum_{s \in \{\emptyset, 1995, \cdots, 2003\}} Prob_i^{enter}(s) \hat{x}_{it}(s).$$

7. Check the market clearing condition as

$$\sum_i \hat{b}_{it}(\mathbf{P}) + \bar{B}_t^{fringe}(P_t) = 0 \ \forall t = 1995, \cdots, 2003.$$

8. Repeat steps 1-7 until the market clearing conditions are satisfied.

In practice, I stop the iteration when the following condition is satisfied:

$$\max_{t=1995,\cdots,2003} \left| \sum_i \hat{b}_{it}(\mathbf{P}) + \bar{B}_t^{fringe}(P_t) \right| < 1000.$$

This criterion is sufficiently tight so that the absolute value of the price change is in the order of magnitude of 1e-1.

To update the price in the above procedure, I construct the following rule that exploits the market clearing conditions and the optimality conditions. Denote a current candidate of an equilibrium price by $\mathbf{P}^l = \{P_t^l\}_{t=1995}^{2003}$. The next candidate of price P_t^{l+1} is given by solving the equation

$$\sum_i \sum_s P_{i,enter}(s) \cdot TC'^{(-1)} \left(\hat{\lambda}_{it}(\mathbf{P}^l, s) - P_t^{l+1} \right) + \bar{B}_t^{fringe}(P_t^{l+1}) = 0,$$

where $\hat{\lambda}_{it}(\mathbf{P}^l, s)$ is the prediction of the shadow values under the current candidate of prices $\mathbf{P}^l$.

B Special Cases of Structural Model

Appendix B introduces the special cases of the structural model I introduced in the paper. Those cases are used in counterfactual simulations.

B.1 Case without Permit Banking and with Transaction Costs

I explain the case in which firms are not allowed to bank emissions permits. Once I shut down permit banking, permit holding h_{it} is no longer a state variable in the model. However, the dynamic consideration still plays a role due to abatement investment and participation decisions.

I first consider individual optimization problems. Consider the case in which a firm is a trader. The problem is given by

$$\begin{aligned} V_{it}^1(I_{it}, R_{it}) = \quad & \max_{\{q_{jt}\}_j, b_t} \quad \pi_{it}\left(\{q_{jt}\}_j\right) - \left(P_t b_{it} + TC(b_{it})\right) + \beta V_{i,t+1}(1, R_{i,t+1}) \\ & \text{s.t.} \quad e_{it}\left(\{q_{jt}, \rho_{jt}\}_j\right) = a_{it} + b_{it}. \end{aligned}$$

Note that the choice of $\{q_{jt}\}_j$ and b_t does not affect the continuation value. The optimality conditions of the problem are given by equation (3.6) and (3.7).

Next, consider the case in which a firm is a non-trader:

$$\begin{aligned} V_{it}^0(I_{it}, R_{it}) = \quad & \max_{\{q_{jt}\}_j, b_t} \quad \pi_{it}\left(\{q_{jt}\}_j\right) + \beta V_{i,t+1}(0, R_{i,t+1}) \\ & \text{s.t.} \quad e_{it}\left(\{q_{jt}, \rho_{jt}\}_j\right) \leq a_{it}. \end{aligned}$$

In this case, a firm may not use all the permits, due to the capacity constraints of production. The emissions level is given by

$$e_{it}^* = \min\left\{a_{it}, e_{it}^{max}\right\},$$

where e_{it}^{max} is the emissions level when a firm is facing zero shadow costs of permits.

Other components, including the participation and the investment decisions are the same as in the baseline case (i.e., the case that includes both permit banking and transaction costs).

B.2 Shutting Down Transaction Costs without Permit Banking

This section explains the case in which I shut down both transaction costs and permit banking. In this case, I do not need to consider the participation decision.

Given an emissions rate R_{it}, the individual problem in period t is given by

$$\begin{aligned} \Pi_{it}(R_{it}) = \quad & \max_{\{q_{jt}\}_{j \in J_{it}}, b_{it}} \quad \pi_{it}\left(\{q_{jt}\}_j\right) - P_t b_{it} \\ & \text{s.t.} \quad e_{it}\left(\{q_{jt}, \rho_{jt}\}_j\right) \leq a_{it} + b_{it}. \end{aligned}$$

The FOC is by $\partial \pi_{it} / \partial q_{jt} = P_t \; \forall j$. This gives the optimal choice for $e_{it}(R_{it}, P_t)$ and $b_{it}(R_{it}, P_t)$.

The investment decision for Phase II is then given as

$$\begin{aligned} W_{i,2000}(R_i^1) = \quad & \max_{R_i^2} \sum_{t=2000}^{2003} \beta^{t-2000} \Pi_{it}(R_i^1) - \Gamma(R_i^2, R_i^1) \\ & s.t. \quad R_i^2 \leq R_i^1 \end{aligned}$$

and the problem for Phase I is

$$\begin{array}{ll} \max\limits_{R_i^1} & \sum\limits_{t=1995}^{1999} \beta^{t-1995}\Pi_{it}(R_i^1) + \beta^{2000-1995}W_{i,2000}(R_i^1) - \Gamma(R_i^1, R_i^0). \\ s.t. & R_i^1 \leq R_i^0. \end{array}$$

To close the model, consider an equilibrium of the permit market. The permit price should satisfy the market clearing conditions:

$$\sum_i b_{it}^*(R_{it}^*(P), P_t) + \bar{B}_t^{fringe}(P_t) = 0 \ \forall t = 1995, \cdots, 2003.$$

B.3 Shutting Down Transaction Costs with Permit Banking

I now consider the case with permit banking. In the absense of transaction costs, Rubin (1996) has shown that the equilibrium path of permit prices grows at the rate of β^{-1} as long as the aggregate banking is positive, which implies that

$$\begin{aligned} P_{t+1} &= \beta^{-1}P_t. \\ \Longleftrightarrow P_t &= \beta^{-(t-1)}P_{1995}. \end{aligned}$$

The optimal decision on emissions, given the emissions rate, is determined by $\partial\pi_{it}/\partial q_{jt} = P_t \ \forall j$, which is the same as the one in appendix B.2. As I discussed in section 3.7.1, individual decisions on net purchase and banking are not determined from the model, because the current shadow value $\lambda_t = P_t$ is equal to the discounted marginal value of banking $\beta\lambda_{t+1} = \beta P_{t+1} = P_t$. In other words, banking and trading decisions are arbitrary as long as a firm can produce the level of emissions determined by the optimality condition.

Now I consider investment decisions. The continuation value at the beginning of Phase II is given by

$$\begin{aligned} V_{i,2000}(h_{i,2000}, R_i^2) &= \sum_{t=2000}^{2003} \beta^{t-2000}\left[\pi_{it}\left(\{q_{jt}\}_j, R_i^2\right) - P_t b_{it}\right] + \beta^{2003-2000}CV(h_{i,T+1}) \\ &= \sum_{t=2000}^{2003} \beta^{t-2000}\left[\pi_{it}\left(\{q_{jt}\}_j, R_i^2\right) - P_t \cdot (e_{it} - a_{it})\right] \\ &\quad + \beta^{2003-2000}\left\{CV(h_{i,T+1}) - P_T h_{i,T+1}\right\} \\ &\quad + \sum_{t=2000}^{2003} \beta^{t-2000}P_t h_{it} + \sum_{t=2000}^{2002} \beta^{t-2000}P_t h_{it+1} \\ &= \sum_{t=2000}^{2003} \beta^{t-2000}\left[\pi_{it}\left(\{q_{jt}\}_j, R_i^2\right) - P_t \cdot (e_{it} - a_{it})\right] \\ &\quad + \beta^{2003-2000}\left\{CV(h_{i,T+1}) - P_T h_{i,T+1}\right\} + P_{2000}h_{i,2000}, \end{aligned}$$

where the last equality uses the equilibrium relationship $\beta P_{t+1} = P_t$. The investment problem is

$$\begin{aligned} W_{i,2000}(h_{i,2000}, R_i^1) = \quad & \max_{R_i^2} \quad V_{2000}(h_{i,2000}, R_i^2) - \Gamma(R_i^2, R^1). \\ & s.t. \quad\quad R_i^2 \leq R_i^1. \end{aligned}$$

Note that $h_{i,2000}$ does not affect the optimal investment level of R_i^2.

The continuation value at the beginning of Phase I is given as

$$\begin{aligned} V_{1995}(h_{i,1995}, R_i^1) \quad = \quad & \sum_{t=1995}^{1999} \beta^{t-1995} \left[\pi_{it}\left(\{q_{jt}\}_j, R_i^1\right) - P_t(e_{it} - a_{it})\right] \\ & + \beta^{1999-1995}\left(\beta W_{2000}(h_{i,2000}, R_i^1) - P_{1999} h_{i,2000}\right). \end{aligned}$$

The investment problem is similar to the one in Phase II.

Finally, I consider the market clearing condition. By aggregating the transition equation of permit holding (3.2) over individual firms and time, we have

$$\sum_{t=1995}^{2003} E_t(P_t) + H_{T+1} = \sum_{t=1995}^{2003} A_t + \sum_{t=1995}^{2003} B_t, \tag{B.1}$$

where $E_t = \sum_i e_{it}(P_t)$, and other uppercase variables are similarly defined. The market clearing condition in each period is

$$B_t + \bar{B}_t^{fringe}(P_t) = 0.$$

By putting this condition into equation (B.1), we have

$$\sum_{t=1995}^{2003} E_t\left(\beta^{-(t-1)} P_{1995}\right) + H_{T+1}(\beta^{-(T-1)} P_{1995}) = \sum_{t=1995}^{2003} A_t + \sum_{t=1995}^{2003} -\bar{B}_t^{fringe}\left(\beta^{-(t-1)} P_{1995}\right).$$

The equilibrium price P_1 is determined by this equation, and thus the whole path of the equilibrium price.

C Derivations and Proof

C.1 Derivation of $\partial EV_t(h_t, I_t)/\partial h_t$

I omit an index i for a particular firm for expositional purposes. I focus on the derivation of $\frac{\partial EV_t(h_t,0)}{\partial h_t}$. Recall that

$$EV_t(h_t, 0) \quad = \quad \int \max\left\{V_t^0(h_t), V_t^1(h_t) - F_t - \epsilon\right\} dG(\epsilon).$$

By the chain rule, we have

$$\frac{dEV_t(h_t,0)}{dh_t} = \frac{\partial EV_t}{\partial V_t^0}\frac{dV_t^0}{dh_t} + \frac{\partial EV_t}{\partial V_t^1}\frac{dV_t^1}{dh_t}.$$

First, we derive $\frac{\partial EV_t}{\partial V_t^k}$ for $k = 0, 1$. This is an application of the Williams-Daly-Zachary theorem (see Theorem 3.1 in Rust, 1994):

$$\frac{\partial EV_t(h_t)}{\partial h_t} = \mathbb{P}(h_t)\left\{P_t + T'(b_t^{trade})\right\} + (1-\mathbb{P}(h_t))\,\pi'(e_t^{not}).$$

By using the interchange of integration and differentiation, we have the following (I omit h_t for expositional purposes in the following derivation):

$$\begin{aligned}
\frac{\partial EV_t}{\partial V_t^1} &= \frac{\partial}{\partial V_t^1}\int \max\left\{V_t^1 - F_t - \epsilon, V_t^0\right\} dG(\epsilon) \\
&= \frac{\partial}{\partial V_t^1}\int_{\Upsilon^1}(V_t^1 - F_t - \epsilon)dG(\epsilon) + \frac{\partial}{\partial V_t^1}\int_{\Upsilon^0} V_t^0 dG(\epsilon) \\
&= \int_{\Upsilon^1}\frac{\partial}{\partial V_t^1}(V_t^{trade} - F_t - \epsilon)dG(\epsilon) + \int_{\Upsilon^0}\frac{\partial}{\partial V_t^1}V_t^0 dG(\epsilon) \\
&= \int_{\Upsilon^1} dG(\epsilon) \\
&= \mathbb{P}_t(h_t),
\end{aligned}$$

where Υ^1 is the set of ϵ such that a firm chooses to participate, i.e., $\Upsilon^1 \equiv \{\epsilon : V_t^1 - F_t - \epsilon > V_t^0\}$, and Υ^0 is similarly defined. Note that we can apply the similar derivation to obtain $\frac{\partial EV_t}{\partial V_t^0} = 1 - \mathbb{P}(h_t)$.

Next, we calculate $\frac{\partial V_t^k}{\partial h_t}$ for $k = 0, 1$. The derivation is a direct application of the envelope theorem (or the Benveniste-Scheinkman formula):

$$\frac{\partial V_t^k}{\partial h_t} = \lambda_t^k,$$

where λ_{it}^k is the Lagrange multipliers in the corresponding optimization problems. Thus, we obtain

$$\frac{dEV_t(h_t,0)}{dh_t} = \mathbb{P}_t(h_t)\lambda_t^1 + (1-\mathbb{P}_t(h_t))\,\lambda_t^0.$$

C.2 Proof of Comparative Statics

This subsection shows the comparative statics in section 3.7. I omit an index of firm i for expositional purposes. I also re-write the profit from the electricity market as a function of emissions volume e_{it} for the purpose of exposition. The optimization problem for the trader is now given by

$$\begin{aligned}
\max_{e_t, b_t, h_{t+1}} \quad & \pi_t(e_t) - (P_t b_t + TC(b_t)) + \beta EV_{t+1}(h_{t+1}, 1) \\
\text{s.t.} \quad & e_t + h_{t+1} = a_t + h_t + b_t, \\
& h_{t+1} \geq 0,
\end{aligned}$$

and the problem for the non-trader is similarly defined. For simplicity, I assume that the non-borrowing constraint is not binding: $\mu_t = 0$ throughout the proof.

First, I focus on the case in which a firm does not participate in trading. I show that e_t and h_{t+1} are increasing in h_t. The optimality condition in this case is given by

$$\pi_t'(e_t) = \beta \frac{\partial EV_{t+1}(\overbrace{a_t + h_t - e_t}^{=h_{t+1}}, 0)}{\partial h_{t+1}}.$$

Using an implicit function theorem, we have

$$\begin{aligned}
\frac{de_t}{dh_t} &= -\frac{-\beta EV_{t+1}''(a_t + h_t - e_t, 0)}{\pi''(e_t) + \beta EV_{t+1}''(a_t + h_t - e_t, 0)} \\
&= \frac{\beta EV_{t+1}''(a_t + h_t - e_t, 0)}{\pi''(e_t) + \beta EV_{t+1}''(a_t + h_t - e_t, 0)} \\
&\in (0, 1),
\end{aligned}$$

where $EV_{t+1}''(\cdot, 0) = \partial^2 EV_{t+1}(h_{t+1}, 0)/\partial h_{t+1}^2$. Note that both $EV_{t+1}(\cdot)$ and $\pi(\cdot)$ are concave functions, so that their second derivatives are non-positive. Because $h_{t+1} = (a_t + h_t) - e_t$,

$$\frac{\partial h_{t+1}}{\partial h_t} = 1 - \frac{\partial e_t}{\partial h_t} > 0.$$

Next, I consider the case in which a firm participates in trading. The optimality conditions, given by equations (3.6) and (3.8), are

$$\begin{aligned}
\pi'(e_t) - P_t - TC'(e_t + h_{t+1} - a_t - h_t) &= 0 \\
\pi'(e_t) - \beta EV_{t+1}'(h_{t+1}, 1) &= 0.
\end{aligned}$$

Taking the total derivative of these equations with respect to h_t, we have

$$\begin{aligned}
(\pi'' - TC'')\frac{\partial e_t}{\partial h_t} + (-TC'')\frac{\partial h_{t+1}}{\partial h_t} + TC'' &= 0 \\
\pi'' \frac{\partial e_t}{\partial h_t} - \beta EV_{t+1}'' \frac{\partial h_{t+1}}{\partial h_t} &= 0.
\end{aligned}$$

Solving these equations gives me

$$\begin{aligned}
\frac{\partial e_t}{\partial h_t} &= \frac{-TC''}{\pi'' - TC'' - TC''\frac{\pi''}{\beta EV''_{t+1}}} > 0 \\
\frac{\partial h_{t+1}}{\partial h_t} &= \frac{\pi''}{\beta EV''_{t+1}} \frac{\partial e_t}{\partial h_t} > 0.
\end{aligned}$$

Thus, h_t increases both e_t and h_{t+1}. Finally, h_t decreases b_t because

$$\begin{aligned}
\frac{\partial b_t}{\partial h_t} &= \frac{\partial e_t}{\partial h_t} + \frac{\partial h_{t+1}}{\partial h_t} - 1 \\
&= \frac{-TC''}{\pi'' - TC'' - TC''\frac{\pi''}{\beta EV''_{t+1}}} + \frac{-TC''\frac{\pi''}{\beta EV''_{t+1}}}{\pi'' - TC'' - TC''\frac{\pi''}{\beta EV''_{t+1}}} - 1 \\
&= \frac{-TC'' - TC''\frac{\pi''}{\beta EV''_{t+1}}}{\pi'' - TC'' - TC''\frac{\pi''}{\beta EV''_{t+1}}} - 1 \\
&= \frac{-\pi''}{\pi'' - TC'' - TC''\frac{\pi''}{\beta EV''_{t+1}}} < 0.
\end{aligned}$$