Does Electricity Restructuring Benefit the Environment? Theory and Evidence from Intertemporal Emission Trading in the U.S. SO_2 Allowance Market

Fan Zhang*

Abstract

Intertemporal trading of emission permits allows for the banking of permits for future use or sale. In this paper, I explore the effects of increased uncertainty over future output prices, input costs and productivity levels on the temporal distribution of emissions. In a dynamic programming setting, the permit price is a convex function of each of these three sources of uncertainty. Increased uncertainty about future market conditions increases the expected permit price and causes risk-neutral firms to reduce *ex ante* emissions to smooth out marginal abatement costs over time. The convexity results from the asymmetric impact of changes in counterfactual emissions on marginal abatement costs. Empirical analysis corroborates the theoretical prediction. I find that increased price volatility induced by electricity market restructuring could explain 7-10% of the allowances banked during Phase I of the U.S. sulfur dioxide trading program. Numerical simulation suggests that high uncertainty may generate substantial initial compliance costs, thereby deterring new entrants and reducing efficiency; sharp emission spikes are also more likely to occur under high uncertainty scenarios. These results are subjected to a number of robustness tests.

*Penn State University. Contact: fxz10@psu.edu. I thank Philippe Ambrosi, Li Gan, William Hogan, Dale Jorgenson, Juan Pablo Montero, Erich Muehlegger, Zmarak Shalizi, Robert Stavins, Tao Wang and David Wise for helpful comments and suggestions. I acknowledge financial support from the World Bank Development Economics Research Group and Heinz Family Foundation. This paper also benefited from discussions with seminar participants at Harvard University, Penn State University, University of Michigan, Brown University, Arizona State University, Camp Resources XIV and the 2006 USAEE meeting. All errors remain solely of the author.

1 Introduction

Cap-and-trade emission permit systems that allow permits to be traded across compliance periods (hereafter *intertemporal emissions trading* or *bankable emission permit trading*) are witnessing growing regulatory interest as a cost-effective way to reduce total emissions. The U.S. sulfur dioxide (SO_2) emission trading program is one of the first, and by far the most extensive application of bankable emission permit trading. Under Title IV, firms are not only allowed to transfer allowances[1] for emissions of SO_2 between facilities, but also to bank them for use in future years. Emission permit trading is also a centerpiece of the Kyoto Protocol, which allows participating nations to trade and bank greenhouse gas permits under the Framework Convention on Climate Change (Intergovernmental Panel on Climate Change, 1996). [2]

Despite the considerable interest in intertemporal emission trading, important theoretical and policy issues surrounding this trading mechanism remain unexplored. Although the theoretical literature on tradable emission permits began a discussion regarding the efficiency and properties of their use as early as 1970s[3], most of the literature considers trading between units, implicitly within a single time period. Theoretical analyses of intertemporal emission trading have only recently appeared.[4] These studies typically assume firms have perfect foresight[5]. Neither the theory nor the empirical assessment of the implications of uncertainty has been examined thoroughly. In this paper, I seek to fill this gap in the literature.

This paper makes two specific contributions. First, I introduce uncertainty into the intertemporal trading model, which is theoretically more interesting and empirically more relevant. In this model, a firm decision regarding permit trading is an *ex ante* choice in the sense that optimal emissions and permit banking decisions depend not only on current output and input prices, but also on expectations of future prices. Assuming risk neutrality and a competitive permit market, I show that a mean-preserving increase in electricity price volatility would decrease *ex ante* emissions. Second, I empirically test the theoretical

[1] For the purpose of this paper, I use the terms 'permits' and 'allowances' interchangeably.

[2] In fact, intertemporal trading goes far beyond the realm of air pollution control. It applies to the trading of goods whose existence is statutorily generated but privately transferable and storable. For instance, this concept applies to the banking and trading of wetland development rights, the allocation of Corporate Average Fuel Economy (CAFE) credits, the temporal and spatial trading of lead in gasoline, and the transferring of federal funds between and within banks to meet reserve requirements.

[3] Examples include Montgomery (1972), Hahn (1984), and see Titenberg (1985) and Cropper and Oates (1992) for thorough reviews.

[4] Studies that analyze intertemporal emission trading include Rubin (1996), Cronnshow and Kruse (1996), Kling and Rubin (1997), Schennach (2000), Yates and Cronshaw (2001), Leiby and Rubin (2001), Stevens and Rose (2002), Sedio and Marland (2003), Maeda (2004), Stranlund, et al. (2005), van Steenbergh (2005), Feng and Zhao (2006), Wirl, F. (2006).

[5] Studies that mention uncertainties include Schennach (2000), Feng and Zhao (2006).

prediction in a real trading program, the U.S. SO_2 allowance trading. To the best of my knowledge, this is the first study that quantitatively estimates the effects of uncertainty on emissions trading based on actual market data. Although the analysis is conducted in the context of the U.S. SO_2 allowance trading program, the model is flexible enough to be extended to other intertemporal trading initiatives, such as the global carbon trading program, for which uncertainty is a prevalent feature in many of the policy parameters.

Based on numerical analyses, I also revisit the often considered regulatory instruments choice between taxes and quotas along the lines of Weitzman (1974). In a multi-period dynamic framework, I find that uncertainty can qualitatively affect the conventional thoughts concerning when it is socially optimal to use quantity-based v.s. price-based environmental policies.

In the first part of the paper, I develop a stochastic dynamic optimization model of risk-neutral price-taking firms, which use high- and low-sulfur coal to produce electricity. Current prices are observed, but future prices evolve stochastically. In equilibrium, the marginal abatement costs of all firms are equalized with the permit price in each period; for each firm, the expected present marginal abatement costs are also equalized across time periods. When firms expect higher compliance costs in the future, they will demand positive permits, bidding up the permit price (and therefore the marginal abatement cost) until current and expected future prices are equal.

The permit price is subject to electricity price shocks, which alter firms' abatement costs by changing the industry-wide counterfactual emissions level[6]. I prove analytically that the permit price, as well as firms' marginal abatement costs, are convex functions of the electricity price. By Jensen's inequality, this leads to a positive relationship between the marginal abatement cost and the increase in volatility of the stochastic electricity price. This convex relationship results from the asymmetric impact of the electricity price change on the change in marginal abatement costs. Because abatement costs are convex in the abatement level (as long as marginal productivity is decreasing), when the electricity price increases (and therefore the abatement level increases), the marginal abatement cost increases faster than it decreases when the electricity price falls. When uncertainty is pronounced, very high and very low electricity prices become more likely, and this asymmetric impact becomes more salient. In the presence of extreme electricity prices, firms would have a much higher incentive to reduce *ex ante* emissions and to shift permits to subsequent periods so as to smooth out the marginal abatement costs over time. This conclusion holds with or without perfect competition in the output market. In addition, I extend the model to allow uncertainties to enter through input costs and industry average productivity. The conclusions are similar, following the same line of reasoning.

I also discuss firms facing an additional rate-of-return regulation. My analysis shows

[6]Counterfactual emissions are the emissions that would prevail when there is no environmental regulation.

that while a competitive firm tends to reduce more emissions than a regulated firm in each time period, as long as regulatory treatments on allowance transactions remain constant over time, one should not expect qualitative differences in banking behavior between regulated and unregulated firms.

In the second part of the paper, I test the theoretical prediction using data on the U.S. SO_2 allowance trading market from 1996 to 2004. Although ten years have passed since the launch of this program, most empirical analyses evaluating its market performance were conducted during the early years of the program (Schmalensee et al., 1998; Ellerman et al., 2000); fewer have focused on the banking dynamics. Indeed, allowances banking by regulated sources, brought on by overcompliance with environmental regulation in Phase I, has been a major component of the trading program, to such an extent that the level of the banked allowances has largely exceeded expectations. I provide empirical evidence that shows how the seemingly irrational banking behavior can be at least partially explained by firms' responses to market uncertainties induced by electricity restructuring. Specifically, a one percent increase in electricity price volatility measured by annualized standard deviation of percentage price change is on average associated with a 0.84% decrease in the annual emission rate. Overall, electricity restructuring may explain 7-10% of the total amount of the banked allowances during Phase I of the SO_2 trading program.

To estimate the impact of market uncertainty on social welfare, I simulate the banking pattern, emission stream and time path of permit price under various degrees of electricity price uncertainties through 2020. Numerical results are suggestive and they are consistent with the econometric estimation of the elasticity of emission rate with respect to price volatility. Specifically, the results indicate that high uncertainty generates large compliance costs in the early years, leading to efficiency loss and deterring new entrants. From an environmental perspective, sharp emission spikes are more likely to occur under market-wide uncertainty shocks or high uncertainty scenarios. When a pollutant creates convex flow damage, the disproportionate distribution of emissions could dramatically increase health hazards. In closing, I discuss implications for climate change policy and the establishment of robust and effective banking regimes.

This paper is related to three strands of literature, one discussing intertemporal emission permit trading, another on capital investment under uncertainty, while the third concerns the impact of electricity restructuring on the environment.

Among previous theoretical investigations of intertemporal permit trading, Schennach's (2000) paper is a first effort to study the implications of uncertainty on the time-series properties of emissions trading. Schennach suggests that the higher the expected electricity price, the lower the emissions in the *ex ante* period. Schennach also emphasizes the role of the non-negativity constraint, a special feature of the U.S. SO_2 allowance trading program, arguing that the expectation of a potential stockout of the allowance account may induce reduction of emissions in earlier periods.[7] Feng and Zhao (2006) also discuss effects of abatement costs uncertainty and conclude that more permits will be banked when the expected marginal value of permits rises. While constituting important steps toward an understanding of the potential consequences of uncertainty, these papers do not answer the question of how increased output price volatility would modify the path of emissions. After all, it is the significant variation, not the level of prices that defines a volatile market.

In spirit, this paper is closer to those of Hartman (1972) and Abel (1983, 1985). In these models, the existence of adjustment costs, together with a constant-returns-to-scale technology, make the marginal revenue product of capital a convex function of output price. Therefore, increased uncertainty about future price increases the expected marginal revenue product for a competitive firm, consequently increasing the intensity of investment. Similarly, I demonstrate the positive relationship between electricity price uncertainty and the expected marginal value of an allowance by applying Jensen's inequality for a convex function. However, the analysis of the marginal value of a permit has no direct analogue in the capital investment literature. In addition, I derive the model in a more general framework, without assumptions of constant-returns-to-scale or perfect competition in the output market.

This paper also contributes to the policy discussion on the implications of electricity restructuring for the environment. Policy debates on the potential environmental impact of restructuring, in large measure, have focused on the effects of market liberalization on the mix of generation technologies (electricity produced from gas, coal, hydro, nuclear and non-hydro renewable sources of energy).[8] I address the question from a new perspective by analyzing the impact of electricity market restructuring on the environmental performance of the single most polluting type of generation technology, coal-burning power generation. I show that, in the short term, electricity restructuring contributed to coal power industry emission reductions by providing incentives for early abatement.

The remainder of the paper is organized as follows: Section 2 provides background on the U.S. SO_2 allowance trading program; Section 3 analyzes the impact of electricity restructuring on the allowance market; Section 4 develops a firm model of intertemporal

[7] I show that uncertainty decreases emissions in the *ex ante* period even when non-negativity constraints are not binding.

[8] For example, Palmer and Burtraw (2006) argue that expanded interregional electricity trading will increase the use of older low-cost coal power plants which would in turn lead to an increase in emissions; Holland and Mansur (2004) show that real-time pricing, an anticipated feature of a competitive market, will shift load from peaking to baseload plants. Depending on which type of plants is dirtier, real-time pricing will have different environmental impacts in different regions. Mansur (2005) suggests that changes in air pollution emissions resulting from the exercise of market power will depend solely on the technologies that dominant firms use to withhold output in contrast with the technologies that the competitive fringe uses to meet demand. Other qualitative analyses raise the concern that a cost-conscious marketplace will invest less in renewable energy.

emissions trading and derives the relationship between emissions banking and uncertainty; Section 5 presents the empirical model and the estimation results. Section 6 discusses numerical simulation, welfare analysis and policy implications and Section 7 concludes the paper.

2 The U.S. SO_2 Allowance Trading Program

The U.S. SO_2 allowance trading program, also known as the Acid Rain Program, was established under Title IV of the Clean Air Act Amendments of 1990 (CAAA90). By creating a national clean air market, it was a grand application of a market-based regulatory approach to achieve emission reduction goals. The basic idea behind permit trading is simple. The regulatory agency first sets a cap that limits the total SO_2 emissions by less than half of their 1980 level (from 18.9 million tons in 1980 to 8.9 million tons by 2001). It then divides the quantity up to a number of tradable allowances and allocates them to individual firms based on their historical heat inputs. Each allowance grants the holder the right to emit one ton of SO_2 emissions. Firms that can reduce emissions relatively cheaply may increase their profits by selling extra allowances; while those for whom reducing emissions is expensive can buy extra allowances from the market. The SO_2 allowance trading program institutionalized a couple of innovations in that it not only allows unlimited trading of permits among firms, but also allows permits to be traded over time. So power producers who reduce emissions below the number of allowances they hold may sell allowances to other firms, or bank them for future use. The only limitation the U.S. Environmental Protection Agency (EPA) imposes on the trading program is that firms cannot borrow allowances from their future allocation. At the "true-up" date (usually the end of March), each unit must submit enough allowances to cover its emissions for that year.

Another important feature of this program is that it was phased-in. Phase I, which ran from 1995 through 1999, affected 263 units at 110 mostly coal-burning (and a few oil-fired units) electric utility plants located in 21 eastern and Midwestern states. Most Phase I units had emissions greater than 2.5 pounds of SO_2 per MMBtu[9] and a generating capacity greater than 100 megawatts (MW). Phase II began in the year 2000. It established a permanent cap of 8.95 million per year and affects all existing utility units with an output capacity of 25 MW and larger, and all new utility units.

Figure 1 shows the annual emission cap, aggregated emissions and banked allowances from 1995 to 2004. The temporal dimension is clearly a key component of this trading program. From 1995 to 1999, 11.65 million allowances were banked, which was about 30%

[9]Btu stands for British thermal unit, a unit of energy frequently used to describe the heat value (energy content) of fuels. MMBtu stands for one million Btus.

of the total allowances allocated during Phase I.[10] These extra allowances were produced through reducing emissions below the allowable standard.

Units banked permits primarily because the program was phased-in: an allowance is perceived to be worth more in later years under the stricter cap of the Phase II. As expected, in 2000 firms began drawing down the bank to ease the transition to Phase II. However, the size of the bank generated in Phase I was unexpectedly large. Some argue that banking in this program has been excessive and was economically inefficient (Ellerman, et al., 2000; Smith, et al., 1998). In addition, the draw-down rate at the beginning of Phase II was lower than previously expected (Ellerman and Montero, 2005). In the remainder of the paper, I explore the question of how to interpret this temporal banking trend.

3 Electricity Market Restructuring and Price Volatility

The implementation of Title IV happens to have coincided with electricity restructuring which dramatically changed the way the power industry was structured and regulated over the past decade. In this section, I briefly review the background of electricity restructuring and its effect on electricity price volatility and the emission allowance market.

Traditionally, in the United States, the three parts of electricity supply-generation, transmission, and distribution-were assumed to be a natural monopoly and were operated by a single utility to exploit economies of scale. Under such a model, within a defined geographical area, one or a small number of firms have exclusive rights to serve retail customers. Because utilities operated as monopolies, they have been subject to extensive price regulation by state and federal regulatory agencies.

Throughout the 1970s and 1980s, technology development weakened the notion of the 'natural monopoly'. Rising electricity prices further increased public pressure for correcting the economic inefficiency of the old regulatory system. Together with a shift in regulatory thinking, electricity restructuring proceeded at an ever-accelerating pace at both federal and state levels in the 1990s. The goal of electricity restructuring is to increase competition in the electricity generation sector.[11] In 1996, the Federal Energy Regulatory Commission (FERC) issued Orders 888 and 889 to open up the transmission network to competitive generators of electricity so as to promote wholesale market competition. In addition, 24 states and the District of Columbia have taken steps to introduce competition in retail markets, including encouraging entry by competitive energy producers and allowing customers to choose their own electricity suppliers.[12]

[10]The number of banked allowances does not include allowances sold at public auction each year, nor does it include contributions from substitution units that entered or exited the market in different years.

[11]The transmission and distribution sectors are still considered as a natural monopoly. They were unbundled from the generation segment and remain regulated and noncompetitive.

[12]For a more comprehensive review of the drivers and the process of electricity restructuring in the United

Restructuring of the wholesale and retail markets has been accompanied by significant changes in the pricing of electricity generation. Before restructuring, electricity price is set administratively on the basis of the average production cost. In contrast, competitive generation prices are determined by market forces. Given open access to the transmission system, a number of auction-based regional wholesale markets were established. In these markets, producers submit bids to supply power and the dispatch order is set by the bids. In most cases, the cost of producing the last MW of electricity to meet demand, namely the marginal production cost of the marginal producer, determines the market-clearing price and is paid to all plants that are dispatched. In the restructured retail markets, the retail rate is linked to wholesale prices, as competitive retail sellers compete with utilities to sell electricity to consumers.[13]

Competitive pricing induced a significant variation in the price of electricity, which can be translated into fluctuating demands for coal generation. The principal drivers behind this volatility are shifts in supply and demand. On the supply side, prices based on marginal costs are sensitive to any factors that affect the operating costs of the marginal generators. Since the operating cost of natural gas-fired power plants exceeds those of most other generation technologies, natural gas power plants usually set the market price. When natural gas prices turn out to be higher or lower then expected, competitive generation prices are directly affected. As has been observed, however the natural gas market behaves, the electricity market behaves similarly.

Natural gas price volatility in recent years was mainly driven by factors exogenous to electricity markets. Natural gas prices in the U.S. had been kept artificially low and stable for over 20 years due to excessive supply. When supply and demand reached a balance in the 1990s, gas prices became responsive to ever-changing market situations, such as persistently colder-than-normal temperatures that increased demand for heating fuel, frozen gas wells and pipelines that reduced regional gas production, and multiple hurricane seasons that disrupted supply. Natural gas is a substitute fuel for oil and petroleum so political unrest in key oil producing nations also contributes to natural gas price volatility(Villar, 2006).

On the demand side, record high temperatures in recent years drove up the demand for electricity. As more and more power plants use natural gas to generate electricity, the growing electricity market tightened the demand and supply balance of natural gas and induced natural gas price spikes. Natural gas price fluctuations in turn exacerbated electricity prices.

In a competitive market, the demand for coal generation can be affected by any or all of the factors that impact natural gas and electricity prices. This is illustrated in Figure 2. The upper part of Figure 2 shows the marginal production costs of coal- and natural

States, see Joskow (1997), U.S. Energy Information Agency (EIA) (2000) and EIA (2003).

[13] Retailers either buy electricity from wholesale markets or generate it themselves.

gas-fired power plants.[14]

Given that the fringe supply from hydroelectric and nuclear generation is typically inelastic and would not be able to move prices, the lower part of Figure 2 plots the cumulative supply that only considers fossil-fuel generation units with the residual demand served by that set of units. The vertical axis shows the offer prices of generation units (the dispatch costs), and the horizontal axis indicates the type of fuel used. The demand curve can be traced from the bids submitted by the buyers indicating at which price they are willing to buy electricity. The market is "cleared" at the point where the supply and demand curves intersect. All units to the left of the equilibrium point are economical to dispatch and will profit from high peak prices.

In the base case represented by the solid line, the cheapest coal power plants are dispatched first. Until the electricity price hits point B, the cumulative supply curve is merely the marginal cost of the coal-fired power plants. When demand increases, more costly coal and gas units are added to the generation mix and the market price increases. The market is cleared at B_e.

Suppose there is an exogenous shock in the natural gas market that increases the natural gas price. Because fuel cost is a key component of the marginal cost of a natural gas-fired plant, the supply curve of gas units shifts up, which is denoted by the dashed line Gas'. As a result, an operator would not be willing to run a natural gas-fired power plant if the market clearing price is lower than A. Assuming the coal price remains stable, more coal capacity will be dispatched and coal plants will cover a longer portion of the supply curve, as indicated in the lower graph. In addition, the cost of supplying peak load electricity also increases, which is reflected by a higher market clearing price, A_e.

The result is quite intuitive. As coal is one of the cheapest and most widely available fuels in the United States, coal prices remain quite stable compared to those of natural gas. When natural gas prices rise sharply, coal prices become more competitive. Power producers will shift toward coal either by increasing the capacity factor of incumbent coal units, or by importing cheaper coal-fired power from other areas. When the price increases to a level higher than the long-run average production cost of coal power plants, new entries by coal units will also be triggered.

Higher demand for coal-based generation means higher demand for SO_2 emission allowances. Since the supply for allowances is generally fixed, increased demand will drive up the allowance price, imposing an industry-wide shock in the national allowance market.

[14] The marginal cost curves are upward sloping indicating that more expensive units are called upon to meet increasing demand. Older and inefficient coal-power plants are more expensive to run. In addition, increased coal-fired generation may drive up the permit price that increases the marginal production costs. Natural gas-fired power plants based on different production technology also have different marginal costs: a typical super peak plant based on a gas turbine has a higher marginal cost than a peak plant based on gas-fired combined cycle gas turbine (CCGT).

Such a shock differs from firm-level idiosyncratic shocks[15], which have a lesser impact on price distribution and would not induce entry. Aggregate shock affects the expected path of permit price through new entries or expansion of current ones. Because aggregate shocks affect each unit in the industry, they are not diversifiable through future or spot markets. Hence, facing aggregate volatilities, the industry may reduce its dependence on the spot market through overcompliance and banking. In doing so, firms adjust *ex ante* emissions in anticipation of future demand and future price changes.

4 Modelling Framework

This section contains the basic theory of intertemporal permit trading under uncertainty. I begin by setting up a firm's dynamic optimization problem, then state and prove Proposition 1 and Lemma 1 on the relationship between uncertainty, and banking/emissions. Next, I show how conclusions can be affected by imperfect competition in the electricity and allowance markets, and by returns to scale of production technology. Finally, I discuss the effects of joint constraints from environmental and rate-of-return (ROR) regulations on banking incentives.

4.1 A Firm Model of Intertemporal Permit Trading under Uncertainty

Consider a risk neutral firm that uses adjustable levels of low- and high-sulfur coal to produce electricity. In each time period, the firm decides the electricity output (g_t), chooses the mix of low- and high-sulfur coal (l_t and h_t), and the amount of allowances (x_t) to buy ($x_t > 0$) or sell ($x_t < 0$) to maximize its discounted present profits for a constrained level of emissions. Uncertainty exists in the supply and demand for electricity. Suppose this uncertainty is characterized by electricity prices (P_e), which is a random variable that follows a Markov process. The probability law of P_e is known to all firms. At the start of period-t, the firm observes electricity price (P_{et}), allowance price (P_{at}), the price for low- and high-sulfur coal (P_{lt} and P_{ht}), and the initial endowment of allowances which is the sum of allowances issued by the government in current period (A_t) and the amount of banked allowances carried forward from the previous period (B_t). Table I provides a thorough description of all model parameters used in the paper.

Firms face a dynamic optimization problem because they must choose how many allowances to save for the future before uncertainties over future prices are resolved. Assuming that firms are price takers in all markets, I model individual firm behavior as an intra-firm game. Taking the strategies of other firms as given, each firm picks a strategy in each time period that is optimal from the firm's perspective in that period. The firm's

[15] Such as an unscheduled outage of a nuclear power plant that increases demand on neighboring baseload coal power plants and increases such plants' demands for allowances.

strategy is thus a map from the Markov state $\Lambda_t = \{P_t, A_t, B_t\}$ to choice variables $\{l_t$, h_t, $x_t\}$, where P_t is a price vector, i.e. $P_t = \{P_{et}, P_{at}, P_{lt}, P_{ht}\}$.

I assume that firms employ three compliance strategies: abating emissions through blending with or switching to low-sulfur coal, purchasing allowances in addition to initial allocation, and adjusting output levels. Other capital intensive strategies such as scrubbing, re-powering or permanently retiring a facility are not considered because regulatory, financial and other uncertainties, during a period of industry restructuring, provide firms incentives to avoid capital intensive investment as long as possible. Indeed, firms have preferred fuel switching/blending and allowances purchasing that require less capital investment: 52% of the 263 Phase I units chose fuel-switching/fuel-blending and 32% chose to purchase allowances.[16] While only 17 units during 1995-2004 have installed new scrubbers (EIA, 1997)

Let $V_i(P_t, B_t, A_t)$ denote firm i's value at time t. The firm's maximization problem can be written as:

$$V_i(P_t, B_{it}, A_{it}) \equiv \max_{l_{it},h_{it},x_{it}} \{P_{et}g(l_{it}, h_{it}) - c(l_{it}, h_{it}) - P_{at}x_{it} + \beta E_t[V_i(P_{t+1}, B_{i(t+1)}, A_{i(t+1)})|P_t]\} \tag{1}$$

$$s.t. \quad B_{i(t+1)} = A_{it} + B_{it} - e_{it}(l_{it}, h_{it}) + x_{it} \tag{2}$$

$$B_{i(t+1)} \geq 0 \tag{3}$$

where β is the discount ratio. In a discrete-time setting, $\beta = 1/(1 + r)$, and r is the risk-free interest rate.[17] E[$\cdot$] is the expectations operator. Based on the realized current price (P_t), the firm formulates expectations on its future value. Eq.(2) is the state equation and defines the stock of banked allowances in period t. Eq.(3) corresponds to the non-negativity constraint, i.e. borrowing against future emission reductions is not allowed. For simplicity, I suppress unit index i in section 4.1.

The production function with low- and high-sulfur coal as two distinct inputs is represented by $g(l, h)$, which is assumed to be quasi-concave, increasing in both arguments, homogenous of degree 1, and twice differentiable everywhere.[18] Most of the previous studies assume output as given (such as Rubin [1996], and Arimura [2002]). In a restructured electricity market, an assumption of fixed output seems untenable. In this model, I assume that producers may alter the output level as an option to meet the required emission standard.

[16]Low-sulfur coal has become more often used also because the rail rates for hauling low-sulfur western coal to higher-sulfur coal-fired plants in the Midwest have declined as a result of railroad deregulation.

[17]The assumption that interest rates are risk-free corresponds to the assumption that firms are risk-neutral.

[18]To include labor and capital inputs in the production function is straightforward and yields an almost identical analysis.

Some previous studies assume that low- and high-sulfur coals are perfect substitutes that differ only in the sulfur contents. For example, Arimura (2002) uses a linear function to describe the production technology with low- and high-sulfur coals as the factor inputs. In reality, low- and high-sulfur coal cannot so easily be substituted for each other. Typically, power plants are designed for a particular type of coal, with which initial performance guarantees are met. Deviations of coal properties from the initial design may result in reduced efficiencies, impaired plant performance, or even serious operating problems, and the damage increases as the deviations multiply.[19]

$c(l, h)$ is the cost function. When a firm undertakes production, it incurs costs that can be described in terms of three components: (1) fuel costs, (2) adjustment costs associated with fuel-blending or fuel-switching,[20] and (3) other fixed costs including capital costs for mixing fuel. Once the binary choice determining whether or not to switch/blend fuel has been made, this sunk cost will have no impact on the factor input ratio. Thus I do not explicitly take account of the initial capital cost in this analysis and assume that both low- and high-sulfur coal are used.

For the model to be tractable, I assume that the adjustment costs are continuous and linear in l.[21] Specifically, I combine the variable adjustment cost and the purchasing cost of low-sulfur coal as an augmented cost and represent the cost function as a standard linear one. That is, $c(l, h) = P_h h + P_l l$, where P_l is the sum of both purchasing cost and the

[19] Before the Acid Rain Program most of the Phase I affected power plants were burning high-sulfur coal, so switching to low-sulfur coal or blending high-sulfur coal with low-sulfur coal may have reduced plants' production performance. For example, low-sulfur coal generally has a lower heat value, resulting in a lower heat rate given a consistent firing rate; some lower-sulfur coal has a higher moisture content and may decrease boiler efficiency, because combustion heat must be used to evaporate the water in the coal; low-sulfur coal also produces higher ash, and/or higher sodium and iron content in the ash, which results in a less efficient heat transfer rate. Furthermore, the boiler heat rate could be adversely affected if the boiler redesign is required to accommodate lower sulfur coal. For more details on the impact of blending high-sulfur coal with low-sulfur coal on the productivity performance, refer to International Energy Agency (1993).

[20] Typically, the transition from high-sulfur to low-sulfur coal incurs the following extra operating and maintenance expenses: (i) more aggressive dust suppression and dust collection procedures, and more diligent housekeeping in coal handling areas, since low-sulfur coal, especially coal from the Powder River Basin (PRB) are very dusty;(ii) more extensive fire protection procedures, due to a higher tendency for spontaneous combustion which increases the risk of fire and explosion; (iii) increased drying requirements in the pulverization process, because some low-sulfur coals have a higher moisture content;(iv) increased pulverizer maintenance because low-sulfur coal are more difficult to pulverize, with their Hardgrove Grindability Indices in the range of 40 to 50; (v) costs incurred to reduce other pollutants as burning western lower sulfur coal results in more particular matter (PM) emissions; and,(vi) increased expenses for storage because lower sulfur coal is usually lower in heating value and requires a larger volume of coal to generate the same amount of power. For a detailed discussion on the impact of lower sulfur coal on an individual plant, see Energy Information Administration (EIA) (1994).

[21] As I will show, the abatement cost is strictly convex.

variable adjustment cost of low-sulfur coal. Although low-sulfur coal may be cheaper than high-sulfur coal in certain areas, given the extra adjustment cost incurred, I assume that P_l is strictly less than P_h.

Finally, I denote the emission function as $e(l, h) = \gamma(\delta_l l + \delta_h h) = \mu_l l + \mu_h h$, where δ_l and δ_h are the sulfur contents of low- and high-sulfur coal ($\delta_l < \delta_h$), γ is the conversion rate from sulfur to sulfur dioxide and μ_l and μ_h are the SO_2 content of low- and high-sulfur coal.

To analyze the above constrained stochastic dynamic optimization problem, consider a firm that is in place for two periods $t = 1, 2$. The Kuhn-Tucker necessary conditions for a maximum at (h^*, l^*, x^*, μ^*) yield the following first-order conditions:

$$P_{a1} = \beta E_1[V_{B_2^*}] + \omega^* \tag{4}$$

$$P_{e1} g'_{f_1^*} = c'_{f_1^*} + (\beta E_1[V_{B_2^*}] + \omega^*) e'_{f_1^*} \qquad (f \ = \ l, h) \tag{5}$$

$$A_1 + B_1 + x_1^* - e_1^* \geq 0, \quad \omega^* \geq 0, \quad \omega^*(A_1 + B_1 + x_1^* - e_1^*) = 0 \tag{6}$$

where ω is the Lagrangian multiplier associated with the non-negativity constraint on B_{t+1} described by (3). $\omega > 0$ if and only if the constraint is binding, i.e. $A_1^* + B_1^* + x_1^* - e_1^* > 0$ implies $\omega^* = 0$.

Eq.(4) is the Euler-intertemporal condition. Eq.(5) discloses that producers choose the optimal levels of coal so that coal's marginal value product equals its marginal cost. The marginal cost includes both the direct production cost ($c'_{f_1^*}$) and the opportunity cost of surrendering the option to use allowances in the future ($(\beta E_1[V_{B_2^*}] + \omega^*) e'_{f_1^*}$). Therefore, expectations on the marginal value of a unit of allowance for period-2 ($E_1[V_{B2}]$) affect current emission decisions.[22]

The second-period optimization problem is

$$V_2 = \max_{l_2, h_2, x_2} P_{e2} g(h_2, l_2) - c(h_2, l_2) - P_{a2} x_2 \tag{7}$$

$$s.t. \qquad A_2 + B_2 - e(h_2, l_2) + x_2 = 0 \tag{8}$$

Eq.(8) shows that firms deplete the allowance bank in the terminal period. The solution (l_2^*, h_2^*, x_2^*,λ_2^*) is described by the following first-order conditions:

$$P_{a2} = \lambda_2^* \tag{9}$$

$$P_{e2} g'_{f_2^*} = c'_{f_2^*} + \lambda_2^* e'_{f_2^*} \qquad\qquad (f \ = \ l, h) \tag{10}$$

[22] $g_f^{'} = \partial g / \partial f, c_f^{'} = \partial c / \partial f, e_f^{'} = \partial e / \partial f (f = l, h)$ represent the marginal productivity, marginal production cost, and marginal emission rate of the two types of coal. Hereafter, $'$ represents the calculation of a derivative.

λ_2 can be interpreted as the shadow value of a unit of banked allowance in period-2. Eq.(9) says that firms will buy or sell allowances such that the shadow value of the marginal allowance equals its market price. The optimal input mix is given by Eq.(10).

An important feature of the above optimal solution is that it is independent of the level of banked allowances B_2. The value function in period 2 is only linearly linked to (B_2) through the profit function. Specifically, the value function in Eq.(7) can be written as

$$V_2 = P_{e2} g(h_2^*(P_2), l_2^*(P_2)) - c(h_2^*(P_2), l_2^*(P_2)) - P_{a2}[A_2 + B_2 - e(h_2^*(P_2), l_2^*(P_2))] \quad (11)$$

where $P_2 = \{P_{e2}, P_{l2}, P_{h2}, P_{a2}\}$. Differentiating Eq.(11) with respect to B_2 gives us the marginal revenue product of allowances:

$$V_{B_2^*} = P_{a2} \quad (12)$$

Substituting this expression for V_{B2} into Eq.(4) leads to a non-arbitrage pricing formula:

$$P_{a1} = \beta E[P_{a2}] + \omega^* \quad (13)$$

The right side of Eq.(13) is the expected return of holding one unit of allowance. It consists of two components: expected present allowance price in period 2 and a convenience yield[23] ω^*. The left side of the equation represents the opportunity cost of carrying an additional unit of allowance, which is an instantaneous gain from selling it in the spot market. Given the substantial number of allowances banked by the industry during Phase I, and that the SO_2 allowance market has been fairly liquid, I assume the convenience yield related to the scarcity of the allowance bank is not a factor affecting banking decisions of individual units, i.e. $\omega^* = 0$.

Combining Eqs.(5) and (12) yields the following policy function for intertemporal emission trading:

$$\beta E[P_{a2}] = \frac{\xi_1^* P_{l1} - P_{h1}}{\gamma(\delta_h - \xi_1^* \delta_l)} \quad (14)$$

where $\xi_1^* = g'_{h_1^*}/g'_{l_1^*}$ is the ratio of the marginal productivities of high- and low-sulfur coal.[24] The right side of Eq.(14) is the additional cost an operator has to pay in order to reduce one ton of SO_2 emissions. It reflects both price and productivity differences between low- and high-sulfur coal. Following Montgomery (1972), emission abatement costs are defined as the difference between unconstrained profits and profits in which the firm adopts an emission level lower than the unconstrained emission level. Therefore, the right side of Eq.(14) presents a notation for marginal abatement cost.

[23] Convenience yield is a concept in finance. It refers to the benefit or premium associated with holding an underlying product or physical good, rather than the contract or derivative product.

[24] The expected permit prices are positive, implying $\delta_h/\delta_l > g'_h/g'_l$.

Eq.(14) together with Eq.(13) exhibit the spatial and temporal efficiency properties of a tradable emission permit regime: in each period, the marginal abatement costs are equalized across firms through the current allowance price (thereby the total pollution reduction cost is minimized)[25]; the present value of the marginal abatement costs are equalized across time periods in an expectation sense. Thus, expectations about higher future allowance prices raise the current abatement level.

Plugging $\beta = 1/(1+r)$ into Eq.(13), we obtain Hotelling's rule under uncertainty.

$$\frac{E(P_{a2}) - P_{a1}}{P_{a1}} = r \tag{15}$$

Eqs.(13) and (14) show that firms have incentives to save allowances for future use (forward banking) every time they expect the discounted future allowance price to be greater than the current market price; at the industry level, such forward banking will drive up the current allowance prices, as well as the current marginal abatement costs, to reflect the expectation of future allowance prices. Eq.(15) shows that forward banking is to prevent the expected allowance price from increasing at a rate higher than the interest rate.

4.2 Uncertainty, Banking and Emission

Although a price is given for each individual unit in the allowance market, allowance price is endogenously determined by the aggregate behavior of the generating units. Previous theoretical analysis of emission permit trading reveals that when allowed to trade with one another in a competitive allowance market, units will collectively behave like a central planner who efficiently allocates emission permits to each unit in a manner that minimizes total costs (Rubin, 1996; Schennach, 2000; Feng and Zhao, 2006). This suggests a model of aggregate industrial behavior as that of a single representative unit, and to solve the equivalent problem without considering internal spatial trading. For simplicity of exposition, I assume the representative agent produces electricity according to the Cobb-Douglas production function[26] $g(l, h) = Gl^{\alpha}h^{1-\alpha}$, where G is a productivity parameter, and $0 < \alpha < 1$ is the share of low-sulfur coal. To avoid confusing increasing price volatility with increasing price trends, I consider electricity price P_e evolves following a mean-preserving stochastic process with the mean equal to $\overline{P}_e$. Formally, I define the probability distribution function of P_e as $f(\cdot, \theta)$ such that

$$\int P_{e2} df(\cdot, \theta) = \overline{P}_e \quad \forall \theta \tag{16}$$

[25]This conclusion is based on the assumption that firms have interior solutions, i.e. both low- and high-sulfur coal are used. If firms only use one type of coal, marginal abatement costs are not equalized between firms having interior solutions and firms having corner solutions; however, an emission trading program still yields a cost effective result.

[26]In Appendix A, I extend the model to a more general CES production function and prove that the conclusions do not change.

where θ is an index of the mean-preserving spread and if $\theta' > \theta, f(\cdot,\theta)$ second-order stochastically dominates $f(\cdot,\theta')$ (or $f(\cdot,\theta')$ is more risky than $f(\cdot,\theta)$). Therefore, the value of θ characterizes the level of market-wide risk. The representative firm's optimization problem in period 2 is simplified by leaving out the term x:

$$\max V = P_{e2} g(l_2, h_2) - c(l_2, h_2) \tag{17}$$

$$s.t. \quad A_2 + B_2 = e(l_2, h_2) = \mu_l l_2 + \mu_h h_2 \tag{18}$$

There is no closed-form solution for the above optimization problem. Nonetheless, I prove analytically in Appendix A that the marginal profitability of allowances $\partial V / \partial B$, or the allowance price P_a, is convex in the stochastic variable P_e. This leads to a negative relationship between *ex ante* emissions and the level of uncertainties about electricity prices.

Proposition 1. *Increasing uncertainty over electricity price generates lower ex ante emissions and higher banking in the following sense: For $\theta' > \theta, B(\theta') > B(\theta)$, and $e_i(\theta') < e_i(\theta)$, where θ is an index of the mean-preserving spread of electricity price, B is the industry's total banked emissions permits, and e_i is the individual unit's ex ante emissions.*

Proof. Because the marginal profitability of allowances is convex with respect to P_e, it follows directly from Jensen's inequality that an increase in the mean preserving spread of P_e increases the expected marginal value of allowances. According to Eq.(14), in anticipation of higher future marginal value of allowances, firms will reduce *ex ante* emissions by increasing current marginal abatement costs, leading to an increased aggregate stock of allowances at the industry level. □

It is essential that the marginal value of allowances be convex with respect to electricity prices to derive the above conclusion. This convexity reveals an asymmetric distribution of future marginal values of allowances due to output prices changes. To understand the intuition, note that because the total number of allowances is fixed, and is less than the emissions expected to be produced by all of the affected units, the rise of electricity prices increases the counterfactual emissions (through the mechanism explained in section 3), as well as the total required pollution reduction. Since abatement costs are convex (further discussion of this property appears in the next section), marginal abatement cost rises with the quantity of abatement. Therefore, when electricity price increases, the marginal abatement cost increases faster than it decreases when electricity price falls. So the potential gain from saving an additional unit of allowance when electricity price increases is higher than the potential loss when electricity price decreases. When uncertainty is more pronounced, very high and very low electricity prices are more likely, and this asymmetric relationship becomes more salient. In the presence of extreme prices, firms would have a higher incentive to save allowances as the potential gain is much higher than the potential loss.

In addition, across multiple time periods, the convexity effect also works through a firm's ability to vary the input of allowances in response to the resolution of uncertainty. When a 'bad' shock occurs, such that the stock of allowances exceeds the desired stock of allowances, firms can choose not to use extra allowances. Thus, the expected profit from saving a unit of allowance today equals $\mathrm{E}[\max(\beta P_{a2} - P_{a1}, 0)]$. The gain from a 'good' shock is unchecked, while the loss from a 'bad' shock is bounded below. A unit of allowance is like a set of American call options on future production, which is worth more when good and bad outcomes are more extreme (with the same expected mean value).

Based on a similar analysis, I show in Appendix B that the marginal value of allowances is also a convex function of input costs: the prices of low- and high-sulfur coal. The intuition follows the same line: the fluctuation of input costs changes the counterfactual emission levels, leading to an asymmetric probability distribution for the marginal profit of allowances: negative shocks which increase the input costs will reduce the marginal value of allowances less than positive shocks will increase them. Hence, when there is a mean preserving spread in the distribution of P_l or P_h, which raises the expected marginal profitability of allowances, the industry will save more permits in equilibrium.

Lemma 1. *The greater the uncertainty in input costs P_l and P_h, the lower the ex ante emissions.*

Proof. See Appendix B. □

Finally, note that the industry average productivity (G) enters the profit function in exactly the same way as the electricity price (P_e). Therefore, introducing uncertainty in productivity given the price is the same as introducing uncertainty in a price given productivity. Hence, the marginal value of allowances is also convex in industry-wide productivity. However, there is a conceptual difference between these two types of uncertainties. Price is usually subject to both upward and downward shocks, while industry productivity, usually affected by technology development, generally moves in one direction. That is, technological innovation would increase rather than decrease productivity. Although, it is inappropriate to analyze the effect of uncertain technological progress within the previous framework, it follows directly from the above analysis,[27] that in anticipation of new technology which may improve the productivity of coal-fired power plants, firms will bank more allowances. The intuition is: as coal-fired power plants become more competitive, the demand for coal-produced electricity will increase and, therefore, the expected marginal value of allowances also increases.

Before moving on, it is important to note that assuming a perfectly competitive allowance market, the above conclusions are derived regardless of a firm's market positions (net seller or buyer) and the initial allocation of allowances.

[27] Also see Eq.(42) in Appendix A

4.3 Imperfect Competition and Returns to Scale

Proposition 1 and Lemma 1 are proved under the assumptions that individual firms are price-takers and that production technology is linearly homogeneous. In this section, I discuss the roles of perfect competition and returns to scale. I show that imperfect competition in an electricity market and decreasing returns to scale do not affect the negative relationship between uncertainty and emissions. However, this negative relationship may not be robust given increasing returns to scale or imperfect competition in the allowance market.

4.3.1 Imperfect Electricity Market

To isolate the impact of imperfect competition in the electricity market, assume the technology is still described by a homogenous of degree one Cobb-Douglas production function: $g = Gl^{\alpha}h^{1-\alpha}$. Suppose an individual firm faces an isoelastic demand curve $P_e = g^{(1-\varphi)/\varphi}W$, where $\varphi(\varphi \geq 1)$ is a markup coefficient with $\varphi = 1$ corresponding to perfect competition. W is an exogenous stochastic demand shifter that captures industry-wide shocks. Under these conditions, the value function is equal to:

$$V = W(Gl^{\alpha}h^{1-\alpha})^{\varepsilon} - P_l l - P_h h - P_a x \tag{19}$$

where $\varepsilon = 1/\varphi$, $\varepsilon \leq 1$, ε can be considered as a returns-to-scale parameter. Appendix C shows that with a value function described by Eq.(19), the convexity of the marginal value of allowances for output price and input costs persists. Therefore, the fact that an electricity market may not be perfectly competitive does not affect the conclusion regarding the negative relationship between uncertainty and *ex ante* emissions given constant returns to scale, perfect competition in allowance market, as well as risk neutrality.

This result confirms the intuition accompanying the foregoing analysis. Recall that a crucial insight from Eq.(11) is that abatement decisions do not depend on either past or future allowance stocks. This lack of 'intertemporal links' holds true as long as firms are price takers in the allowance market, and it does not depend on the elasticity of the demand curve facing individual firms in the electricity market. Therefore, an industry-wide shock will similarly impact abatement decisions for a competitive firm and a monopolist with constant returns to scale in the electricity market. How many allowances are saved now affects profits in the future, but not the level of emissions in the future. As such, any increase in the expected marginal profitability of allowances, including the one caused by an increase in market uncertainty, raises the emissions banked today.

One concern about the foregoing analysis is that the industry itself may face a downward-sloping demand curve even when individual firms are perfectly competitive and have constant returns to scale. When price is endogenously determined by the industry's output, the amount a price can rise under good industry-wide outcomes is limited by the entry of new firms and the expansion of existing ones. If investment is irreversible, as shown by Pindyck (1993), there is no similar mechanism to prevent price from falling under bad demand outcomes. As a result, a mean-preserving distribution of future output prices might not be sustained. However, in the power market, since fossil-fired peak load power plants, which are usually the marginal producers, can be relatively costlessly and temporarily turned on or shut down in response to the realized market price, the possibility of a mean-decreasing distribution of future output price is reduced. So the operational flexibility of peak load power plants weakens the notion of irreversibility which is instrumental in Pindyck(1993)'s analysis.

4.3.2 Returns to Scale

Analysis so far assumes constant returns to scale. Relaxing this assumption, however, does not introduce any additional difficulty. Eq.(19) reveals that a decrease in returns to scale operates exactly like an increase in the markup coefficient and vice versa. Therefore, conclusions of Proposition 1 and Lemma 1 carry over entirely to the case of decreasing returns to scale.

This result has an intuitive interpretation. Note that Proposition 1 and Lemma 1 hinge on the convexity of abatement costs. In principle, this convexity arises from decreasing marginal productivity of factor inputs.[28] To see this, observe from Eq.(14) that the marginal abatement cost is positively related to the ratio of marginal products of high- and low-sulfur coal (ξ). In the case of declining marginal productivity, when producers use a greater share of low-sulfur coal to reduce emissions, ξ increases as the marginal product of low-sulfur coal declines relative to that of high-sulfur coal. Therefore, abatement costs are convex in that a lower level of emissions is associated with higher abatement costs at the margin. In obtaining this result, diminishing marginal productivity is the paramount factor. Because decreasing returns to scale (as well as constant returns to scale) guarantee diminishing marginal productivity (given that the production function is quasi-concave)[29], the negative relationship between emissions and uncertainty extends to the case of decreasing returns to scale (either resulting from imperfect competition, or diseconomies of scale technology, or both). In fact, some cases of increasing returns may also satisfy diminishing marginal productivity. For example, suppose $\varepsilon > 1$, but $0 < \alpha\varepsilon < 1$; the marginal productivity of l decreases with the increase of l. However, as the returns to scale (ε) become larger and larger, given the insights gleaned here, the inverse relationship between emissions and uncertainty would eventually lose its strength.

[28]Increasing marginal abatement costs or decreasing marginal productivity both imply that the firm attains a regular minimum in solving the problem.

[29]A formal mathematical proof is available upon request.

4.3.3 Imperfect Allowance Market

So far I have assumed a competitive allowance market, where the distribution of future allowance price is independent of an individual firm's abatement decisions. The effects of uncertainty are mediated through the equilibrium behavior of all firms and the resulting impact on prices of an allowance market. A logical question to explore is what results if at least one of the units exercises market power in the allowance market. In this case, the current emission of the dominant firm would affect the expected path of the marginal value of allowances. For a dominant firm, Eq.(13) becomes

$$\frac{\partial P_{a1}}{\partial x_1} x_1 + P_{a1} = \beta E[\frac{\partial P_{a2}}{\partial x_2} x_2 + P_{a2}] + \omega^* \tag{20}$$

The price function P_a and its relationship with P_e is intractable without further assumption of the price- or quantity-setting behavior of the dominant firm. However, since the dominant firm has the flexibility to make *ex post* decisions on x_2 after the value of the random demand for coal-produced electricity is known, the permit price P_a becomes less convex with respect to P_e. When the ability of the dominant firm to affect the permit price increases, the firm would respond less and less to changes in the level of uncertainty. Qualitatively, I show that imperfect competition in an allowance market threatens the negative relationship between price uncertainty and *ex ante* emissions. Further analysis is needed based upon more detailed assumptions about the structure of the allowance market and the strategic behavior of a dominant firm.

5 Empirical Analysis

Building on previous discussions, this section empirically explores electric utilities' responses regarding emissions reduction to price fluctuations in the U.S. electricity markets. The analysis is based on a panel dataset consisting of 208 Phase I coal-fired generating units from 1996 to 2004. Before deploying the empirical specification, it is worth commenting on the basic structure of a coal-burning power plant. A steam-electric coal power plant consists of three main components: boiler(s), generator(s), and stack(s). Coal is burned in a boiler(s) to generate steam, which is transmitted to a generator(s) where electricity is produced. During the coal combustion process, SO_2 is produced and disposed of through an exhaust stack. A flue-gas desulphurization system (also called a scrubber) can be installed in the stack to reduce SO_2 emissions. In this paper, I refer to the complex of a boiler and the generator(s) connected to it as a generating unit. Because environmental regulation is implemented at the boiler level, following Arimura (2002), Carson, et al.(2000), Montero (1999), and Ellerman and Montero (1998), estimations in this paper are conducted at the level of generating units.

5.1 Econometric Specification

Assume a generating unit i has a production function of the following form: $g = Gl^a h^b$ ($a > 0, b > 0$). Multiplying factor inputs l and h on both sides of equation (5) gives the input demand functions for l and h:

$$l_t = \frac{P_{et} g_t}{P_{lt} + (\beta E_t[P_{a(t+1)}] + \omega)\mu_l}, \qquad h_t = \frac{P_{et} g_t}{P_{ht} + (\beta E_t[P_{a(t+1)}] + \omega)\mu_h} \tag{21}$$

Substitute l and h from (21) into the emissions function, divide both sides by g_t, and take logs yield the following expression for units' emission rate:

$$ln(\frac{e}{g})_t = lnP_{et} + ln[\frac{\mu_l}{P_{lt} + [\beta E_t(P_{a(t+1)}) + \omega]\mu_l} + \frac{\mu_h}{P_{ht} + [\beta E_t(P_{a(t+1)}) + \omega]\mu_h}] \tag{22}$$

Equation (22) shows that the emission rate of a unit at time t is related to the prevailing electricity price P_{et}, factor input prices (P_{lt} and P_{ht}), and the expected future allowance price $E_t[P_{a(t+1)}]$. As postulated by the theoretical analysis, $E_t[P_{a(t+1)}]$ is positively correlated with the variance of P_{et}. The probability of a potential stockout of allowances measured by ω and other unit specific characteristics may also affect emission performance. Therefore, the emission rate of a unit i is estimated by the following reduced form function:

$$\begin{aligned} Y_{it} \;=\; & \beta_0 + \beta_1 \triangle P_{eit} + \beta_2 lnP_{eit} + \beta_3 lnP_{at} + \beta_4 lnP_{lit} + \beta_5 lnP_{hit} + \beta_6 Z_{it} + \beta_7 R_{it} \\ & + \beta_8 Z_{it} \triangle P_{eit} + \beta_9 R_{it} \triangle P_{eit} + \sum_{j=1}^{20} \nu_j S_j + \sum_{t=1997}^{2004} \kappa_t T_t + \alpha_i + u_{it} \end{aligned} \tag{23}$$

where the dependent variable $Y_{it} = ln(\frac{e}{g})_{it}$ is the observed annual average SO_2 emission rate (in log form) of unit i in year t. Emission rate is calculated by dividing the total annual emissions (tons) by the annual electricity output in megawatt hours (MWh).

$\triangle P_{eit}$ is the electricity price volatility, the key variable of interest in this analysis. Monthly fuel purchasing database reveals that plants purchase coal from spot markets throughout the year. Therefore, decisions regarding how much of each type of coal to buy and how many allowances to hold can be adjusted in response to monthly electricity price fluctuations. I measure $\triangle P_{eit}$ as the standard deviation of the percentage change (between two adjacent months) in monthly average electricity price to industrial customers in year t in the state where unit i is located. [30] The calculation is described by the following equation:

$$\triangle P_{eit} = \sqrt{E\{[\frac{P_{eitm}}{P_{eit(m-1)}} - E(\frac{P_{eitm}}{P_{eit(m-1)}})]^2\}}$$

[30] Industrial prices are the most volatile and least protected by PUC regulation. On a monthly scale, the volatility of industrial prices is driven primarily by changes in fuel costs. However, because of the existence of long-term contracts, using industrial prices may underestimate actual price volatility in the spot market.

where $E[\cdot]$ represents the mean of the sample. m indexes the month. As overall underlying prices tend to be higher in more volatile periods, this definition facilitates normalization and reasonable comparisons across high and low price levels. The coefficient of $\triangle P_{eit}$ provides a measure of the elasticity of annual average emission rate to electricity price volatility. A negative coefficient will provide supporting evidence for the theoretical prediction.[31]

P_{eit} is year t annual average retail electricity price to industrial customers at the state where unit i is located. Given input costs, increased electricity prices will increase the marginal value of allowances. Therefore, the coefficient associated with P_{eit} is expected to be negative. Current allowance price (P_{at}) is an indicator of the market's perception of future prices. A higher P_{at} would reduce current individual units' emissions and the coefficient of P_{at} would likely be negative. The coefficient associated with low-sulfur coal price (P_{lit}) is expected to be positive, while the sign of the coefficient of high-sulfur coal price (P_{hit}) is ambiguous.[32]

Z_{it} is a vector of unit specific characteristics that includes: $SCRUBBER_{it}$, a dummy constructed to be 1 if a scrubber is installed.[33] AGE_{it}, is calculated by subtracting the year of initial boiler operation from the calendar year. I estimate the model allowing AGE_{it} to enter with a quadratic specification. $HEATRATE_{it}$ is a measure of unit efficiency in transferring energy into electricity. It is calculated by dividing the Btu content of the fuel input by the net megawatt hours of power output .[34] $LnCAP_i$ is the log of the boiler nameplate capacity in megawatts. The variable is included to capture possible economies of scale. As previous literature (Avallone and Baumeister, 1996; Joskow and Schmalensee, 1987) suggests that the advantage of larger size tends to deteriorate as scale becomes very large, I therefore include the first and second order terms of $LnCAP_{it}$ in the regression. $WORKLOAD_{it}$ is the ratio between the actual operating hours during year t and the maximum working hours of a year (8640 hours).[35] $INITIAL_{it}$ is the initial

[31]The above analysis implicitly assumes current price fluctuation as a proxy for expected price uncertainty in the future. One concern with this specification is whether current price uncertainty reflects plant operators' expectations of future price uncertainty at the time of making operation decisions. To evaluate whether historical price uncertainty provides insights into expectation of future price changes, I assume managers perfectly predict price volatility in the future ($\triangle P_{ei(t+1)}$), and test for the response of current emission rate to future price fluctuation. This alternative specification does not change the result qualitatively.

[32]Holding the output price constant, a change in input fuel prices has two substitution effects: the substitution between the two types of coal, and the substitution between low-sulfur coal and allowances. When low-sulfur coal prices increase, both substitution effects raise emission rates. However, when the high-sulfur coal price increases, the two substitution effects work in opposite ways, leaving the sign of P_{hit} indeterminate.

[33]Scrubber can remove up to 90% of SO_2 emitted. The coefficient associated with $SCRUBBER_{it}$ is expected to be significantly negative.

[34]Because of the inverse relationship between heat rate and production efficiency, the sign of the coefficient for $HEATRATE$ is expected to be positive.

[35]It is generally understood that frequent ramping up and down tends to increase the level of emissions.

allocation of allowances (tons) issued by the EPA.[36] $MUNI_{it}$ is a dummy equal to 1 when the generating unit is municipally or cooperatively owned. An extensive literature has discussed the relative inefficiency of publicly owned facilities. If $MUNI$ units are less cost conscious, operational activity of these units will be less responsive to output price fluctuations. Therefore, I also control for the interaction term between $MUNI$ and the measure of electricity price volatility ($\triangle P_{eit}$). The coefficient for the interaction term is expected to be positive.

In addition to environmental regulation, electric utilities in regulated states face an extra rate-of-return (ROR) regulation. To control for the impact of ROR regulation on emissions, I include a regulatory status indicator variable and its interaction term with electricity price volatility ($\triangle P_{eit}$). Specifically, the regulatory variable, $RETAILACESS_{it}$, takes the value 1 if the state where unit i is located has begun retail access to industrial customers during year t, 0 otherwise.

A simple analysis is sufficient to show that the marginal abatement cost of a regulated units is lower than the equilibrium allowance price.[37] The wedge between them is the extra profit a regulated firm is allowed to earn either through a positive return on allowance expenditures or a relaxation on profit constraints. Therefore, I expect the coefficient for $RETAILACCESS$ to be negative. I have no priors regarding the interaction term. On one hand, regulated units may be less cost conscious to allowance expenses and be less responsive to price volatility; on the other hand, in anticipation of restructuring, regulated units may have higher incentive to accumulate permits since moving towards restructuring implies the eventual loss of cost recovery.

State dummies S_j are included to control for time-invariant, state-specific environmental regulatory policies.[38] Year dummies T_t picks up year-specific differences in emission performance, such as secular technology or productivity shocks common to all units. α_i is assumed to be an unobservable time-invariant unit-specific characteristic and u_{it} measures idiosyncratic shocks to units' operating performance. $\beta_0, \beta_1, \ldots, \beta_9$, ν_j and κ_t are the coefficients. [39]

I construct the variable $WORKLOAD$ to capture the impact of different operating practices between base load and peak load plants on emissions. The coefficient of $WORKLOAD$ is expected to be negative.

[36] Although emission decisions of a price-taking unit are generally independent of its allowances endowment, $INITIAL$ could be inversely related to the convenience yield (ω) when non-negativity constraints are binding. However, since the non-negativity constraints are inconsequential during Phase I of the trading program, I expect estimation on $INITIAL$ to be positive, but the effect is likely to be tenuous.

[37] The analysis is available upon request.

[38] For example, many generating units were subject to sulfur restrictions contained in State Implementation Plans which were enforced prior to the Acid Rain Program and are still in effect. Some of these local regulations, such as those in Kansas, Michigan, Wisconsin, New York and New Hampshire, are more stringent than those of the Acid Rain Program and therefore consistently affect the emission rates of the units in those states.

[39] Although theoretical models in previous sections do not solve for x, and therefore the theory remains

5.2 Disturbance Term Structure and Robustness Checks

One potential concern for regression model (23) is that if dirty coal power plants passed on high environmental control costs to the consumers, electricity price may be endogenous. To account for the possibility of endogenous electricity pricing, I employ an instrumental variable approach, using the annual average natural gas price delivered to electric utilities (Png (electric power price)) (including transportation cost and taxes etc.) as an instrument for electricity prices in that same state. Natural gas prices are generally determined by weather conditions (hot summer and cold winter days), the prices for substituting fuels (oil and gasoline), and industrial economic activities[40], but plausibly not endogenous to the environmental performance of coal-fired power plants.

A major motivation for electricity restructuring was to remedy the problem of high electricity prices in the Northeast and California. If less clean and less efficient coal power plants in the Northeast contributed to the region's high electricity prices, which in turn induced electricity restructuring, then the regulatory variable may also be endogenous to emission rates. Given that restructuring decisions were generally made before 2000, I conducted reduced sample estimation based on observations from 2000 to 2004 and found that the results do not change.

The equilibrium allowance price could be endogenously determined by units that abuse market power. Emission price endogeneity may be particularly relevant during the first few years of the trading program when the market was not liquid enough and the price determination process might have involved significant interplay of supply and demand between only a few market players. I use the annual average natural gas wellhead price (Png (wellhead price)) as an instrument for the allowance price.[41] Natural gas and coal are competing fuels for electricity generation. Fluctuations in natural gas prices have the potential to influence the market share of coal-based generation, which is a major factor driving allowance prices. I also drop the potential noisy observations in years 1995 and 1996 and compare the results with those from full sample estimation.

---

silent on the relationship between unit-level banking and uncertainty (when the current market price reflects the discounted expected future allowance price, a unit would be indifferent between banking and selling/purchasing allowances). Nonetheless, based on equation (23) and controlling for electricity output, I still test whether the percentage change in banked allowances of unit i between two time periods has any bearing on output price uncertainty (the dependent variable is $[B_{it} - B_{i(t-1)}]/B_{i(t-1)})$, where B_{it} is the amount of banked allowances by unit i in year t). In this regression, the coefficient of electricity price uncertainty ($\triangle P_{eit}$) is statistically significantly positive, suggesting that a one percent increase in electricity price volatility is on average associated with a 2.4 to 3.4 percent increase in the size of the banked allowances at the unit level. For conciseness, I do not report the regression here.

[40] Natural gas is the most versatile fuel. It is used as home heating and industrial fuels, and is increasingly used as a vehicle fuel.

[41] Wellhead price is the value of natural gas at the mouth of the well. In general, the wellhead price is considered to be the sales price obtainable from a third party in an arm's length transaction.

The endogeneity of low- and high-sulfur coal costs may arise as an issue if coal prices are determined by the demand for coal of some individual units. Even when electric generating units are price-takers in coal markets, because coal is a differentiated product, both in terms of quality (sulfur content, Btu content, moisture content etc.)[42] and location, coal suppliers (producers and carriers) may have great latitude in formulating prices.

Following Ellerman and Montero (1998), I use low-sulfur coal price premium ($LSPremium$) and the distance of a unit from the Powder River Basin ($DPRB$) as proxies for high and low-sulfur coal prices available to the units. I control for the distance with a third degree polynomial to reflect a nonlinear relationship between coal prices and the unit's location in relation to Powder River Basin (PRB) coal.[43]

As another robustness check, I exclude 2004 data from the analysis. In 2004, the spot price of SO_2 allowances, which had been steady at about 100-200 dollars/ton since 1995, was increased three-fold after the EPA proposed the Clean Air Interstate Rule, a policy which will lower the SO_2 emission cap by two-thirds beginning in 2010. Although this price response provides further evidence on units' intertemporal optimization behavior, it may swamp the effect of price volatility on emissions trading.

Finally, the results could also be affected by sample attrition issues. Starting in 1998, many plants were divested to non-utilities and dropped from the sample because cost data of nonutility generating facilities are not reported to the public. If divestiture decisions were driven by unobservable unit specific characteristics, for example, if divestitured plants tend to be more competitive and produce fewer emissions given a level of output, the estimation would be biased. To assess if divestiture creates an attrition bias, I constrain the estimation to a balanced sub-panel containing units that remain in the database through 2004. Because the sample selection problem would be the most severe in this specification if observations were not missing at random, comparing results from the balanced sample estimation with those from the unbalanced sample provides a test on the significance of the potential sample attrition problem.

The stochastic disturbance (u_i) in the estimation equations are assumed to be correlated across observations.[44] To obtain robust standard errors, I adjusted standard errors for

[42]Some fuel contracts specify more than a dozen attributes of coal qualities.

[43]A unit's location in relation to PRB coal will to some extent reflect the actual cost of low-sulfur coal and other competing coals. This is because PRB produces the most of the low-sulfur coal in the U.S., with the cheapest coal-mine prices, while transportation costs factor importantly into the delivered price of coal. In fact, for some western coal hauls, transportation costs account for up to 75% of delivered fuel costs. (EIA, 1995). However, the distance to PRB coal does not affect low-sulfur coal price uniformly. Low-sulfur coal from central Appalachia becomes more competitive for units 1000 miles away from PRB coal. It is expected that low-sulfur coal price will first increase with the increase of distance to PRB coal, then decrease as more alternative low-sulfur coal become available.

[44]In the fixed effects estimation of emission rate, the estimated average first-order autocorrelation coefficient is 0.26.

clustering by unit in the following estimations.

5.3 Data and Estimation Results

I began construction of the dataset with all privately and publicly owned Phase I coal-fired generating units. For these units, I built a panel dataset beginning in 1996, the first year for which coal consumption data are available, and ending in 2004, the last year for which the allowances transaction data were updated. The data are collected and merged from several data sources to obtain information concerning annual aggregate productions, quality and quantity of coal used, annual SO_2 emissions, allowances allocated and banked, electricity and fuel prices, regulatory statuses, as well as unit specific characteristics. This merging process reduced the sample size, both because of differences in units covered by various datasets, and because divestitures removed plants from the reporting database after 1999. The final dataset is unbalanced and composed of 208 Phase I coal-fired generating units. All prices are adjusted to real terms using a 5 percent discount rate and presented in 1995 dollars. Details of the dataset collection and construction procedures are provided in Appendix D.

Table 1 presents summary statistics. Table 2 offers the unit-year observations on the number of units affected by electricity restructuring, those which installed scrubbers, switched to low-sulfur coal, or used only high-sulfur coal for production. Table 2 shows that the number of scrubbers installed remains almost constant (except for sample attrition that reduces the number of the observations), suggesting that changes in emission rates may not be primarily caused by the installation of scrubbers.

Columns (1) and (2) of Table 3 report results for estimating equation (23) via unit fixed and random effects models, with $ln(\frac{e}{g})_{it}$ as the dependent variable. The estimated coefficients of $\triangle P_{eit}$ are similar, both negatively and statistically significant from zero. This result is consistent with the theoretical prediction of a negative correlation between emission rate and electricity price volatility. Based on the random effects specification, a one percent increase in electricity price volatility is associated with a decrease in units' annual average emission rate by 0.84 percent. This means a one-standard deviation increase in electricity price volatility would induce a sample average unit to reduce annual emissions by 677 tons.

Columns (3)-(6) of Table 3 explore robustness to alternative specifications or sample. Column (3) presents instrumental variables (IV) fixed effects estimation, using natural gas electric power price and natural gas wellhead price as instruments for electricity and SO_2 prices. I use Generalized Methods of Moments (GMM) to obtain consistent and efficient estimates. The Cragg-Donald F statistic is 56.71 and is significant at 5 percent level, indicating that the instruments are relevant and the model is identified. The Hansen J test statistic is 5.711 and the overidentifying restriction is not rejected at any reasonable level.

The other test statistics are reported in the notes for Table 3. The IV estimation results closely resemble those in columns (1) and (2), suggesting that the endogeneity of electricity and allowances prices does not seem to be a particular concern.

Columns (4) - (6) are based on GLS random effects estimation. Column (4) restricts samples to 1997 - 2003; column (5) reports results from a balanced panel that contains only units that remain in the sample through 2004; while the regression controlling for unit locations in relation to PRB coal and the observed difference between low- and high-sulfur coal prices is reported in column (6). Overall, these alternative specifications do not qualitatively change coefficient estimation for electricity price volatility in either magnitude or statistical significance.

Estimations on the other explanatory variables are generally consistent with prior expectations. The estimates on *SCRUBBER* show that the installation of a scrubber would on average reduce emission rates by about 90%. The coefficient on *HEATRATE* suggests that less efficient production is associated with higher emissions. Higher allowances prices, as well as higher electricity prices, induce lower emission rates. In contrast, there is a positive relationship between low- and high-sulfur coal prices and emission rates. However, coefficients of these price variables are not statistically distinguishable from zero in all models. In addition, the three distance coefficients are each statistically significant. The reversed U-shaped profile derived from the distance coefficients echoes the results from Ellerman and Montero (1998), suggesting that low-sulfur coal prices first increase then fall with the units' distance to PRB coal mine. Finally, note that units that have had not experienced retail competition are more responsive to price volatility, plausibly due to expectations on upcoming unfavorable change in cost recovery rules.

I use actual price volatility level in 1992 ($\triangle \mathrm{Pe}_{1992}$) as what would have prevailed in the absence of electricity restructuring and compute the corresponding emission rate based on results for model (2) in Table 3. With expected counterfactual price volatility, extra allowances - the difference between counterfactual emissions and actual emissions is about 10-percent of the total banked allowances of the sample units. Following the same exercise but using the actual price volatility in 1995 or the average price volatility level during the pre-restructuring period (1990-1995) as the counterfactual, the increased price volatility can explain about 7-percent of the total allowances banked by the sample units during Phase I.

6 Welfare Analysis and Policy Implications

To gain further insights into the effects of uncertainty on the time-series behavior of banking and emissions, in this section, I numerically simulate the banking pattern, emission stream and permit price path over time resulting from varied price volatilities and compare the findings to results that would occur in the absence of uncertainty. Given different social damage functions, I analyze the potential welfare impact of increased price volatility and discuss the policy implications.

6.1 Simulation of Emission Trading under Different Price Volatilities

In the following, I analyze emissions banking at the industry level and ignore internal spatial trading of permits. Firms are assumed to collectively maximize the aggregate present value of profits over a planning horizon spanning from 1995 to 2020, based on a production function $g = G(l^{\alpha}h^{1-\alpha})^{\varepsilon}$. There is no rate-of-return regulation and the allowance market is perfectly competitive. The optimization problem is described by Eqs.(17) and (18) in section 4.2. In addition, I assume the electricity price P_{et} evolves following a mean-preserving stochastic process:

$$P_{et} = \begin{cases} \overline{P}_e + \theta & \text{with probability } q \\ \overline{P}_e & \text{with probability 1-2}q \\ \overline{P}_e - \theta & \text{with probability } q \end{cases}$$

where $\overline{P}_e$ is the expected mean of the electricity price; q denotes the probability that a price moves up or down by θ. Both q and θ measure the magnitude of uncertainty. To be consistent with previous analyses, I vary the value of θ from 0 to 1, while keeping q constant at 0.3.[45] To focus attention on the impact of uncertainty, I maintain a constant realized electricity price in each period at $\overline{P}_e = 4.8$ cents/KWh. Annual initial allocation is 7 million tons in the first five years and is permanently capped at 3.5 million tons from the year 2000. Production parameters are chosen with the following values: $G = 55, a = 0.6, \varepsilon = 0.9$.[46] Discount ratio β is assumed to be 0.95. Values of the following parameters are chosen around the sample means: $\overline{P}_e = 4.8$cents/KWh, u_l=1.64lb/MMBtu, $u_h = 4.41$ lb/MMBtu, $P_l = 120$ cents/MMBtu, $P_h = 100$ cents/MMBtu.[47] Permit price is endogenously determined by the aggregate operational behavior of firms according to Eq.(68). The process of simulation is discussed in detail in Appendix E.

Numerical analysis produces robust patterns in responses of banking and emission to the increase in output price uncertainty. Figure 4 depicts the total amount of banked emission permits as a function of time. The dashed line corresponds to a scenario in which the price

[45] I assess the sensitivity of the results with respect to the value of q.

[46] Ideally, I would estimate the production function based on actual data. However, besides observations on purchasing choices of low- and high-sulfur coal, there are no data on actual inputs of low- and high-sulfur coal. Production factor G is chosen to be large enough so that the emission standard imposes a binding constraint on the production decision. As a sensitivity analysis, I analyze the change in G on the results and find it does not change their qualitative pattern.

[47] The price premium of low-sulfur coal, considering the coal blending adjustment cost, is chosen at 20 cents/MMBtu. I examine the importance of the value on the results in the sensitivity analysis.

spread, θ=0.2, and the emission cap remains constant at 7 million tons across all periods. The shaded, fuzzy line tracks the actual allowance storage through 2004. The other lines indicate a two-stage schedule of declining emission standards with different assumptions about price volatility (θ= 0, 0.2, 0.4, 0.6, 0.8, and 1).

Comparing the dashed line with the other solid lines, it is apparent that when price volatility is low, tightening environmental standards provide a major incentive for producers to reduce emissions below the standards in the early years in order to accumulate credits that can be used when standards become more stringent. When price volatility becomes significant (θ increases), the asymmetric impact of uncertainty plays a more important role in determining the optimal size of the bank. As anticipated, a price spread of $\theta = 1$ generates the largest number of banked permits during Phase I and a zero price volatility generates the fewest. Uncertainty also affects the draw down rate of the bank. When there is no uncertainty ($\theta = 0$), the bank is depleted by the year 2008. When $\theta = 1$, the banking period is significantly extended: the bank continues to grow until 2011. Figure 4 also indicates that the electric industry has been successful in planning emissions banking. The actual banking path follows the optimal routes closely.

Of particular interest to the environmental authority is the effect of uncertainty on the emissions stream. The annual emission flows and the cumulative emissions in each year under different scenarios of price volatility are plotted in Figures 5 and 6. As noted, the emission stream associated with different price volatilities deviate substantially in the beginning and terminal periods. When prices are volatile, it is optimal to emit less in early periods and more in later periods. When these uncertainties are high enough, excessive emissions are observed towards the end of the planning horizon.

Also noted from Figures 5 and 6 is that the size of early emission reductions is significant. Based on the numerical analysis, when $\theta = 1$, producers would have reduced emissions by 48% more than when $\theta = 0$, during the first 5 years of the program; with $\theta = 0.2$ or 0.4, early emission reductions are at the level of 5% and 14%, respectively, which are in the same order of magnitude as the elasticity estimated in previous econometric analysis.

Figure 7 contains the permit price path under different price uncertainties. Permit price equals the marginal abatement cost. Therefore, Figure 7 corresponds to Figure 5 in the sense that the higher the emission level, the lower the permit price. For the high variance cases, permit prices are extremely high initially and plummet at the end. It is worth mentioning that the expected price differs from the realized price. Producers bank emission credits in order to equilibrate expected present value price across compliance periods. However, when uncertainty is extremely high, banking does little to smooth the actual price series.

The simulation model assumes specific parameter values for the production function and input and output prices. To test the sensitivity of results to these assumptions, a number of simulations were run based on different values of $q, G, \alpha, \varepsilon, P_l$, and P_h. The corresponding simulation results on total banked permits are shown in Table 8. The results indicate that the qualitative conclusions do not hinge on the specific parameter values chosen. In addition, as shown in the first two rows of Table 8, the size of the bank is much larger when the probability of price movement, q, increases from 0.3 to 0.4, confirming the relationship between uncertainty and emissions banking from another perspective.

6.2 Welfare Analysis

From the standpoint of economic efficiency, uncertainty shifts emissions abatement to earlier periods, raising abatement costs due to the discounting effect. Assume $\beta = 0.95$, when θ increases from 0 to 1, the sum of the discounted net payoff is reduced by 9%; when $\beta = 0.85$, the net payoff declines by 20%. Furthermore, high initial compliance costs generated by high uncertainty would deter new entrants and have a negative impact on the development of the competitive output market and the emerging environmental market. Although it is generally believed that intertemporal trading creates compliance flexibility that reduces abatement costs and increases efficiency, uncertainty may dampen the cost saving properties of emissions banking.

In addition to cost considerations, depending on the nature of the pollutants, early abatement also has different and important environmental implications. If the pollutants, such as greenhouse gases, create stock damage, voluntary early reduction would yield significant environmental benefits. However, in a finite planning horizon, early abatement increases the degree to which firms will concentrate emissions in later time periods, raising the potential for emission spikes. If the pollutants create flow damages, and if the damage function is convex, emission spikes may even trigger the threshold effect. Emission spikes could also be associated with market structure change. I simulate an exemplary situation in which the price spread (θ) remains at 0.4 from 1995 to 2002, then dropped to 0.2 through 2020. As shown in Figure 8, a sharp emission spike occurs one year before the expected decline in price volatility.

6.3 Policy Implications

It is generally concluded that uncertainty about the cost of controlling carbon dioxide emissions makes price instruments preferable to quantity instruments, because the cost of limiting one ton of emissions is expected to rise as the abatement increases; meanwhile the expected benefit of each ton of carbon reduced is roughly constant because climate change is driven by stock effects rather than flow effects.(Hoel and Karp, 2001; Pizer, 2002).[48] However, for multi-period emissions control, when marginal abatement costs are

[48] The conclusion follows from Weitzman (1974) that when the slope of the marginal cost function is greater than the slope of the marginal benefit function, price instruments are preferable to quantity instruments

also uncertain for regulated sources, a tradable quota system that allows banking creates incentive for early abatement and generates substantially greater environmental benefits than a tax schedule. In addition, since the initial caps on carbon emissions are likely to be relatively undemanding, the expectation of later, more stringent caps will produce even higher reduction in initial years when the cap is non-binding.[49]

On the other hand, if the marginal benefits of abatement are steep when compared to the marginal costs, a quantity instrument without restrictions on the temporal transfer of emissions, may not always be preferable to a price regulation. This is because a quota system exposes firms to volatile market prices, which induces reallocation of emissions in response to observed uncertainty. When marginal damaging effects increase rapidly along with the increase of emission flows, a price instrument would be advisable to directly control the marginal social cost. Another potential solution is to employ a hybrid approach that combines a tradable quota system with safety measures such as restricting the intertemporal trading ratio and/or applying discount to banked permits. The government may also consider incorporating multiple polluting industries into a national trading program so that uncertainties facing one industry can be diversified, and the importance of building up a bank to buffer unexpected price strikes may be reduced.

7 Summary and Conclusion

This paper has extended the existing literature by incorporating uncertainty over the demand for outputs, the supply of inputs, and over technological progress, into the analysis of multi-period emissions trading. Uncertainty affects optimal abatement decisions through its impact on the distribution of future permit prices. Under the assumptions of a competitive permit market and quasi-concave production function, I have shown that there is a convex relationship between the permit price and the different sources of uncertainty. Applying Jensen's inequality discloses that higher uncertainty over stochastic prices and productivity raises the expected value of permits. Since a risk neutral firm that maximizes the sum of discounted profits will always reduce emissions until marginal abatement costs equal the expected permit price (conditional on the existence of an interior solution), firms will emit less in volatile markets than they would if future market conditions were known; consequently, the industry as a whole will accumulate permits at a higher level in an *ex ante* period.

Intuitively, these results are primarily driven by the non-linearity in abatement costs.

(and vice versa), because they are much more likely to minimize the adverse consequences of choosing the wrong level of control.

[49]Currently, the transfer of unused allowances from 2005 - 2007 to the first commitment period under the Kyoto Protocol, i.e. 2008-2012, is not allowed under an EU-wide ban on banking, which, from an environmental point of view, seems a troubling decision.

By changing counterfactual emissions, the fluctuations in the output prices, production costs or productivity, can be interpreted as uncertainty about industry-wide abatement levels. When the marginal cost of reducing emissions rises, in accord with greater reductions of pollutants, uncertainty imposes an asymmetric impact on future permit prices. This occurs because the marginal cost increases faster from a given increase in abatement levels than it decreases from a corresponding decline in abatement. Hence, the higher the volatility, the more likely that firms will gain rather than lose by saving. An immediate implication from the analysis is that the convexity of the cost curve is a key factor in determining the responsiveness of firms' compliance behaviors when confronting uncertainty. Firms may prefer to reduce emissions more in early periods, depending on how steeply the marginal cost rises.

When the basic model is extended to account for imperfect competition in output markets, and decreasing-returns-to-scale technology, the mechanisms through which uncertainty affects firm-level emissions abatement decisions and their impact on industry-wide permits banking remain the same; and therefore, the fundamental results do not change. The conclusions are also not affected by rate-of-return regulation, individual firms' market positions or a firm's initial bank size. This is because the uncertainty considered in the model is at the industry level, it affects market equilibrium price and firm behavior through industry-wide capacity expansion and new entries.

Building on the foregoing analysis, this paper has suggested an explanation for the puzzle of persistent overcompliance with the Acid Rain Program in Phase I. A panel data analysis has revealed that increased price volatility induced by electricity market restructuring could have contributed to 7-10% of the extra emission reductions during Phase I of the SO_2 trading program. From this perspective, electricity restructuring has contributed to emissions reduction in the short-term by providing incentives for early abatement. However, in the long term whether electricity restructuring benefits the environment still depends on whether the incentive is sustainable and whether regulated sources would concentrate emissions during short periods in later years. [50]

Based on numerical simulations, I solved the dynamic optimization problem for firms with a planning horizon from 1995 to 2020. Examining banking and emission trajectories under a variety of price uncertainties, there is consistent time-series evidence that the optimal level of emissions is reduced by an increase of the mean-preserving spread in the electricity price.

Results of these analyses have important policy implications. By showing that the timing of emissions is sensitive to the volatility of the economic environment, I demon-

[50] Also any benefits from early abatement should be compared with the potential cost of price uncertainty. For example, economic and regulatory uncertainties induced by electricity restructuring may have caused the delay of scrubber installation

strate that the environmental impact of uncertainty depends on the degree to which social damages can be assumed to be linear or convex. Therefore, with regard to multi-period emissions control, regulatory policies should take into account both the dynamic effects of uncertainty and the characteristics of the pollutant. The resultant desirability of regulatory intruments could be quite different from that of the traditional view, which originates in Weitzmans (1974) seminal paper. In fact, for a pollutant that creates stock damage (i.e. when the marginal benefit function is relatively flat), a quantity-based banking system encourages early voluntary pollution reduction and could generate sizable environmental gains. Conversely, to control pollutants producing convex flow damage, the steeper the marginal benefit function, the more a tax schedule that directly controls marginal social cost is justified.

It should be noted that this paper does not consider the correlation between multiple sources of uncertainty. An interesting direction for future research would be to analyze the more complicated effects that arise when uncertainties about output prices and input costs are negatively or positively correlated. In a more complete model, one might also take into account uncertainties associated with emissions caps and interest rates.

newpage

References

[1] **Abel,A.B.** "Optimal Investment Under Uncertainty." *American Economic Review*, March 1983, 73(1), pp.228-33.

[2] **Abel,A.B.** "A Stochastic Model of Investment, Marginal q and the Market Value of the Firm." *International Economic Review*, June 1985, 26(2), pp.305-22.

[3] **Averch, H. and Johnson, L.L.** "Behavior of the Firm Under Regulatory Constraint." *American Economic Review*, 1962, 52, pp.1052-69.

[4] **Arimura, T.H.** "An Empirical Study of the SO2 Allowance Market: Effects of PUC Regulations." *Journal of Environmental Economics and Management*, 2002, 44, pp.271-89.

[5] **Baumeister, T. et al**. *Mark's Standard Handbook for Mechanical Engineers.* 8th edn. McGraw-Hill, New York 1980.

[6] **Carlson,C, Burtraw, D, Cropper, M. and Palmer K.** "Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?" *Journal of Political Economy*, 2000, 108(6), pp.1292-1326.

[7] **Cropper, M. and Oates, W.** "Environmental Econocmis: A Survey." *Journal of Economics Literature*, 1992, 30, pp. 675-740.

[8] **Cronshaw, M. and Kruse,J.B.** "Regulated Firms in Pollution Permit Markets with Banking." *Journal of Regulatory Economics*, 1996, 9, pp.179-89.

[9] **Ellerman, A. D. and Montero,J-P.** "The Declining Trend in Sulfur Dioxide Emissions: Implications for Allowance Prices." *Journal of Environmental Economics and Management*, 1998, 36(1), pp.26-45.

[10] **Ellerman, A.D., Joskow, P.L. et al**. *Markets for Clean Air: The U.S. Acid Rain Program*, Cambridge University Press. 2000,

[11] **Ellerman A.D. and Montero,J-P.** "The Efficiency and Robustness of Allowances Banking in the U.S. Acid Rain Program." Working Paper 05-005, MIT Center for Energy and Environmental Policy Research, 2005.

[12] **EMF Report 17**, *Prices and Emissions in a restructured electricity market* Vol I., Standford University, 2001.

[13] **U.S. Energy Information Administration.** "Electric Utility Phase I Acid Rain Compliance Strategies for the Clean Air Act Amendments of 1990." DOE/EIA-0582, 1994, Washington, DC.

[14] **U.S. Energy Information Administration.** "Energy Policy Act Transportation Rate Study: Interim Report on Coal Transportation." DOE/EIA-0597, 1995, Washington, DC.

[15] **U.S. Energy Information Administration.** "The Effects of Title IV of the Clean Air Act Amendments of 1990 on Electric Utilities: An Update." Energy Information Administration Report DOE/EIA-0582(97), Washington, DC, 1997.

[16] **U.S. Energy Information Administration.**"The Changing Structure of the Electric Power Industry 2000: An Update." DOE/EIA-052(00), Washington, DC, 2000.

[17] **U.S.Energy Information Administration.** "Status of State Electricity Industry Restructuring Activity, Washington, DC, 2003.

[18] **Feng, H. and Zhao, J.** "Alternative intertemporal permit trading regimes with stochastic abatement costs." *Resources and Energy Economics*, 2006, 28, pp.24-40.

[19] **Hahn, R. W.** "Market Power and Transferable Property Rights." *Quarterly Journal of Economics*, 1984, 99(4) pp.753-765.

[20] **Hartman, R.** "The Effects of Price and Cost Uncertainty on Investment." *Journal of Economic Theory*, October 1972, 5(2), pp.258-66.

[21] **Hoel, M. and Karp, L.** "Taxes and Quotas for a Stock Pollutant with Multiplicative Uncertainty." *Journal of Public Economics*, 2001, 82, pp.91-114.

[22] **Holland, S.P. and Mansur, E.T.**. "Is Real-time Pricing Green? The Environmental Impacts of Electricity Demand Variance." Working Paper, CSEM WP 136, 2004, University of California Energy Institute.

[23] **Ineternational Energy Agency Fuel Research**. "Fuel Specifications - Impact on Power Station Performance." IEACR/52, London, England, 1993.

[24] **Intergovernmental Panel on Climate Change**(IPCC). *Climate Change 1995. Economic and Social Dimensions of Climate Change: Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel Change.* In J. Bruce, H.P. Lee, and E.F. Haites, eds. Cambridge University Press, Cambridge, 1996.

[25] **Joskow, P. L. and Schmalensee, R.** "The Performance of Coal-Burning Electric Generating Units in the United States: 1960-1980." *Journal of Applied Econometrics*, April 1987, 2(2), pp.85-109.

[26] **Joskow, P.L.**"Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector." *The Journal of Economic Perspectives*, 1997, 11(3), pp.119-38.

[27] **Kling, C. and Rubin, J.** "Bankable Permits for the control of environmental pollution." *Journal of Public Economics*, 1997, 64, pp.101-115.

[28] **Leiby, P. and Rubin, J.** "Intertemporal Permit Trading for the Control of Greenhouse Gas Emissions." *Environmental & Resource Economics*, 2001, 19(3), pp. 229-256.

[29] **Lile, R. and Burtraw, D.** "Guidance for Compliance in the Early Years of the SO_2 Emission Allowance Trading Program." Resources for the Future Discussion Paper 98-35, May 1998.

[30] **Maeda, A.** "Impact of Banking and Forward Contracts on Tradable Permit Markets." *Environmental Economics and Policy Studies*, 2004, 6(2), pp.81-102.

[31] **Mansur, E.T.** "Environmental Regulation in Oligopoly Markets: A study of electricity restructuring", Working Paper, 2005, Yale University.

[32] **Montero, J-P.** "Voluntary Compliance with Market-Based Environmental Policy: Evidence from the U.S. Acid Rain Program." *Journal of Political Economy*, 1999, 107(5), pp.998-1033.

[33] **Montgomery,W.D.** "Markets in licenses and efficient pollution control programs." *Journal of Economics Theory*, 1972, 5, pp.395-418.

[34] **Palmer, K. and Burtraw, D.** "The Environmental Impacts of Electricity Restructuring: Looking Back and Looking Forward." *Environment & Energy Law & Policy Journal*, 2006, forthcoming.

[35] **Pindyck, R.S.** "A Note on Competitive Investment Under Uncertainty." *American Economic Review*, March 1993, 83(1), pp.273-77

[36] **Pizer, W.A.** "Combining Price and Quantity Controls to Mitigate Global Climate Change." *Journal of Public Economics*, 2002, 85, pp.409-434

[37] **Rubin, J.D.** "A model of intertemporal emission trading, banking and borrowing." *Journal of Environmental Economics and Management*, November 1996, 31(3), pp.269-86

[38] **Smith, A.E., Platt, J. and Ellerman, A.D.** "The Costs of Reducing Electric Utility SO_2 Emissions may not be as low as You Might Think." Working Paper 98-010, MIT Center for Energy and Environment Policy Research, 1998.

[39] **Schennach,S.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,pp.189-210.

[40] **Schmalensee, R., Joskow, P.L. et al.** "An Interim Evaluation of Sulfur Dioxide Emissions trading." *Journal of Economic Perspectives*, 1998, 12(3),pp.53-68.

[41] **Sedjo, R.A. and Marland G.** "Inter-trading permanent emissions credits and rented temporary carbon emissions offsets: some issues and alternatives." *Climate Policy*, 2003, 3(4),pp.435-44.

[42] **Stevens, B. and Rose, A.** "A Dynamic Analysis of the Marketable Permits Approach to Global Warming Policy: A Comparison of Spatial and Temporal Flexibility." *Journal of Environmental Economics and Management*, 2002, 44(1), pp.45-69.

[43] **Stranlund, J.K. Costello, C., et al.** "Enforcing Emissions Trading when Emissions Permits are Bankable." *Journal of Regulatory Economics*, 2005, 28(2), pp.181-204.

[44] **Tietenberg, T.H.**, *Emissions trading: an exercise in reforming pollution policy*, Resources for the future, Washington D.C. 1985.

[45] **Van Steenberghe, V.** "Carbon Dioxide Abatement Costs and Permit Price: Exploring the Impact of Banking and the Role of Future Commitments." *Environmental Economics and Policy Studies*, 2005, 7(2), pp.75-107

[46] **Villar, J A. and Joutz, F.L.** "The Relationship Between Crude Oil and Natural Gas Prices." Energy Information Administration, Office of Oil and Gas, October 2006.

[47] **Weitzman, M.** "Prices vs. Quantities." *Review of Economic Studies*, 1974, 41, pp.447-491

[48] **White, M.** "Power Struggles: Explaining deregulatory reforms in electricity markets." Brooking Papers on Economic Activity: Microeconomics. 1996, pp.201-250.

[49] **Wirl, F.** "Consequences of irreversibilties on optimal intertemporal CO2 emission policies under uncertainty." *Resource and Energy Economics*, 2006, 28(2), pp.105-123.

[50] **Yates, A. J. and Cronshaw, M. B.** "Pollution Permit Markets with Intertemporal Trading and Asymmetric Information." *Journal of Environmental Economics and Management*, 2001, 42(1),pp.104-118

Appendices

A Proof of Proposition 1

A.1 Cobb-Douglas Production Function

The optimization problem for the central planner of the industry:

$$\max \quad P_e g(l,h) - c(l,h)$$

$$s.t. \quad A + B = e(l,h) = \gamma(\delta_l l + \delta_h h) = \mu_l l + \mu_h h \tag{24}$$

where $g(l,h) = Gl^{\alpha}h^{1-\alpha}$, $c(l,h) = P_l l + P_h h$, B is the total emissions left at the beginning of the terminal period. Define the Lagrangian expression:

$$L = P_e g(l,h) - c(l,h) + \lambda(A + B - \mu_l l - \mu_h h) \tag{25}$$

The necessary first-order conditions determining a maximum at ($\tilde{l}$,$\tilde{h}$,$\tilde{\lambda}$) are

$$\partial L/\partial B = \tilde{\lambda} \tag{26}$$

$$P_e G\alpha \tilde{l}^{\alpha-1}\tilde{h}^{1-\alpha} - P_l - \tilde{\lambda}\mu_l = 0 \tag{27}$$

$$P_e G(1-\alpha)\tilde{l}^{\alpha}\tilde{h}^{-\alpha} - P_h - \tilde{\lambda}\mu_h = 0 \tag{28}$$

Cross-dividing (27) and (28) results in:

$$\frac{P_l + \tilde{\lambda}\mu_l}{P_h + \tilde{\lambda}\mu_h} = \frac{\alpha}{1-\alpha}\frac{\tilde{h}}{\tilde{l}} \tag{29}$$

The expression on the right side of (29) is the marginal rate of technical substitution (MRTS) between the two types of coal. Eq.(29) says that at the optimum the MRTS between l and h must be equal to their price ratio (including the the opportunity cost of surrendering the option to use allowances in a future period [$\tilde{\lambda}\mu_l$ and $\tilde{\lambda}\mu_h$]).

Define the following: $d_l \equiv (1-\alpha)(P_l + \tilde{\lambda}\mu_l)$ and $d_h \equiv \alpha(P_h + \tilde{\lambda}\mu_h)$ and substitute them into (29)

$$\frac{\tilde{h}}{\tilde{l}} = \frac{(1-\alpha)(P_l + \tilde{\lambda}\mu_l)}{\alpha(P_h + \tilde{\lambda}\mu_h)} = \frac{d_l}{d_h} \tag{30}$$

Solving (30) and (24), we obtain the conditional factor demand functions:

$$\tilde{l} = \frac{d_h}{d_h\mu_l + d_l\mu_h}B, \qquad \tilde{h} = \frac{d_l}{d_h\mu_l + d_l\mu_h}B \tag{31}$$

Substituting $\tilde{l}$ and $\tilde{h}$ from Eq.(31) back into (28) yields:

$$P_e = \frac{1}{G\alpha(1-\alpha)}d_h^{1-\alpha}d_l^{\alpha} \tag{32}$$

Differentiating (32) with respect to $\tilde{\lambda}$:

$$\frac{\partial \tilde{\lambda}}{\partial P_e} = G \frac{d_h^{\alpha} d_l^{1-\alpha}}{\mu_h d_l + \mu_l d_h} > 0 \tag{33}$$

(33) clearly holds for all values of P_e, d_h and d_l.

Differentiating (33) with respect to λ defines the key derivative of the theorem as

$$\frac{\partial^2 \tilde{\lambda}}{\partial P_e^2} = G(\frac{\partial \tilde{\lambda}}{\partial P_e}) \frac{d_h^{\alpha} d_l^{1-\alpha}}{(\mu_h d_l + \mu_l d_h)^2} [\alpha^2 \mu_h \mu_l + (1-\alpha)^2 \mu_h \mu_l + \alpha^2 \mu_h^2 (\frac{d_l}{d_h}) + (1-\alpha)^2 \mu_l^2 (\frac{d_h}{d_l}) - \mu_h \mu_l] \tag{34}$$

Let $d = d_l/d_h$. Note that the minimum value of $\alpha^2 \mu_h^2 d + (1-\alpha)^2 \mu_l^2 (1/d)$ in the last bracketed term of (34) equals $2\alpha(1-\alpha)\mu_h \mu_l$ evaluated at $d = [(1-\alpha)\mu_l]/(\alpha \mu_h)$.

Since by construction, there is $d = [(1-\alpha)(P_l + \tilde{\lambda}\mu_l)]/[\alpha(P_h + \tilde{\lambda}\mu_h)] > [(1-\alpha)\mu_l]/(\alpha \mu_h)$, then $\alpha^2 \mu_h^2 d + (1-\alpha)^2 \mu_l^2 (1/d) > 2\alpha(1-\alpha)\mu_h \mu_l$. Thus $\frac{\partial^2 \tilde{\lambda}}{\partial P_2^e} > 0$, which proves the convexity of the marginal value of allowances with respect to electricity price (recalling that λ is the Lagrangian multiplier and represents the marginal value of a unit of allowances).

A.2 CES Production Function

Now consider a more general CES production function

$$g = G(\alpha l^{\rho} + (1-\alpha) h^{\rho})^{\frac{1}{\rho}} \tag{35}$$

where $\rho < 1$, $\eta = 1/(1-\rho)$ is the elasticity of substitution. α still reflects distribution weight of low-sulfur coal.

The first-order conditions are:

$$P_e G(\alpha \tilde{l}^{\rho} + (1-\alpha)\tilde{h}^{\rho})^{\frac{1}{\rho}-1} \alpha \tilde{l}^{\rho-1} - P_l - \tilde{\lambda}\mu_l = 0 \tag{36}$$

$$P_e G(\alpha \tilde{l}^{\rho} + (1-\alpha)\tilde{h}^{\rho})^{\frac{1}{\rho}-1} (1-\alpha) \tilde{h}^{\rho-1} - P_h - \tilde{\lambda}\mu_h = 0 \tag{37}$$

Cross-dividing (36) and (37), the factor proportion at the optimum is determined by

$$\frac{\tilde{l}}{\tilde{h}} = [\frac{(1-\alpha)(P_l + \tilde{\lambda}\mu_l)}{\alpha(P_h + \tilde{\lambda}\mu_h)}]^{\frac{1}{\rho-1}} \equiv (\frac{d_l}{d_h})^{\frac{1}{\rho-1}} \tag{38}$$

As before, solve $\tilde{l}$ and $\tilde{h}$ from (38) and (24) and substitute them into (36)

$$P_e = \frac{1}{G\alpha(1-\alpha)} [\alpha d_l^{\frac{\rho}{\rho-1}} + (1-\alpha) d_h^{\frac{\rho}{\rho-1}}]^{\frac{\rho-1}{\rho}} \tag{39}$$

Taking partial derivative of (39) with respect to $\tilde{\lambda}$ gives

$$\frac{\partial \tilde{\lambda}}{\partial P_e} = G \frac{[\alpha d_l^{\frac{\rho}{\rho-1}} + (1-\alpha) d_h^{\frac{\rho}{\rho-1}}]^{\frac{1}{\rho}}}{\mu_l d_l^{\frac{1}{\rho-1}} + \mu_h d_h^{\frac{1}{\rho-1}}} > 0 \tag{40}$$

Apparently, (40) holds for all values of P_e, d_h and d_l.

Differentiating (40) with respect to λ gives:

$$\frac{\partial\tilde{\lambda}}{\partial P_e^2} = G\frac{[\alpha d_l^{\frac{\rho}{\rho-1}} + (1-\alpha)d_h^{\frac{\rho}{\rho-1}}]^{\frac{1-\rho}{\rho}}}{(1-\rho)(\mu_l d_l^{\frac{1}{\rho-1}} + \mu_h d_h^{\frac{1}{\rho-1}})^2}(\frac{\partial\tilde{\lambda}}{\partial P_e})T_1 \tag{41}$$

where

$$T_1 \equiv d_l^{\frac{1}{\rho-1}} d_h^{\frac{1}{\rho-1}}[\alpha^2\mu_h^2(\frac{d_l}{d_h}) + (1-\alpha)^2\mu_l^2(\frac{d_h}{d_l}) - 2\alpha(1-\alpha)\mu_l\mu_h] \tag{42}$$

Note that the last bracketed terms of (42) and (34) have a similar structure.

Following the same procedure, we see that because $(P_l + \tilde{\lambda}\mu_l)/(P_h + \tilde{\lambda}\mu_h) > \mu_l/\mu_h$, there is $T_1 > 0$. Hence, $\partial^2\tilde{\lambda}/\partial P_e^2 > 0$. Therefore, given a more general form of production function, the marginal value of allowances is still convex in P_e.

B Proof of Lemma 1

Following the previous discussion, substitute $d_l \equiv (1-\alpha)(P_l + \tilde{\lambda}\mu_l)$ and $d_h \equiv \alpha(P_h + \tilde{\lambda}\mu_h)$ into Eq.(31) to obtain λ as an implicit function of P_l.

$$P_l + \tilde{\lambda}\mu_l = \frac{1}{1-\alpha}[G\alpha(1-\alpha)P_e]^{\frac{1}{\alpha}}\alpha^{\frac{\alpha-1}{\alpha}}(P_h + \tilde{\lambda}\mu_h)^{\frac{\alpha-1}{\alpha}} \tag{43}$$

Taking partial derivative of (43) with respect to P_l and $\tilde{\lambda}$ gives:

$$\frac{\partial\tilde{\lambda}}{\partial P_l} = -\frac{1}{[G(1-\alpha)P_e]^{\frac{1}{\alpha}}(P_h + \lambda\mu_h)^{-\frac{1}{\alpha}}\mu_h + \mu_l} < 0 \tag{44}$$

Differentiate Eq.(44) with respect to λ

$$\frac{\partial^2\tilde{\lambda}}{\partial P_l^2} = -\frac{\frac{1}{\alpha}[G(1-\alpha)P_e]^{\frac{1}{\alpha}}(P_h + \tilde{\lambda}\mu_h)^{-\frac{1}{\alpha}-1}\mu_h^2}{\{[G(1-\alpha)P_e]^{\frac{1}{\alpha}}(P_h + \tilde{\lambda}\mu_h)^{-\frac{1}{\alpha}}\mu_h + \mu_l\}^2}\frac{\partial\tilde{\lambda}}{\partial P_l} > 0 \tag{45}$$

Apparently, (54) holds for all values of P_e, P_l and P_h. So $\tilde{\lambda}$ is also a convex function of P_l. Similarly, we can show that $\tilde{\lambda}$ is convex in P_h. Therefore, increases in the mean preserving spread of P_l and P_h increases the expected marginal value of allowances.

C Proof of Proposition 1 for Imperfect Electricity Market

Consider the following optimization problem for a firm with market power in the electricity market described by Eq.(19) in the paper.

$$V = \max_{l,h,x}[WG^{\varepsilon}l^{\varepsilon\alpha}h^{\varepsilon(1-\alpha)} - P_l l - P_h h - P_a x] \tag{46}$$

$$s.t. \quad A + B = e(l,h) = \gamma(\delta_l l + \delta_h h) = \mu_l l + \mu_h h \tag{47}$$

where $\varepsilon = \frac{1}{\varphi} \leq 1$. W is the stochastic demand shifter and G is the productivity parameter.

The necessary first-order conditions for a maximum at $(\tilde{l},\, \tilde{h},\, \tilde{\lambda})$ are

$$P_a = \lambda \tag{48}$$

$$WG^{\varepsilon}\varepsilon\alpha\tilde{l}^{\varepsilon\alpha-1}\tilde{h}^{\varepsilon(1-\alpha)} - P_l - \tilde{\lambda}\mu_l = 0 \tag{49}$$

$$WG^{\varepsilon}\varepsilon(1-\alpha)\tilde{l}^{\varepsilon\alpha}\tilde{h}^{\varepsilon(1-\alpha)-1} - P_h - \tilde{\lambda}\mu_h = 0 \tag{50}$$

The MRTS is unaffected by the value of ε and is again described by Eq.(29). A similar exercise as in Appendix A provides the following condition where d_h and d_l are functions of $\tilde{\lambda}$.

$$W = \frac{B^{1-\varepsilon}}{G^{\varepsilon}\varepsilon\alpha(1-\alpha)} d_h^{1-\varepsilon\alpha} d_l^{1-\varepsilon(1-\alpha)} (\mu_h d_l + \mu_l d_h)^{\varepsilon-1} \tag{51}$$

Taking partial derivative of (51) with respect to $\tilde{\lambda}$

$$\frac{\partial\tilde{\lambda}}{\partial W} = G^{\varepsilon}\varepsilon\alpha(1-\alpha)B^{\varepsilon-1} d_h^{\varepsilon\alpha} d_l^{\varepsilon(1-\alpha)} (\mu_h d_l + \mu_l d_h)^{2-\varepsilon} T_2^{-1} \tag{52}$$

where

$$T_2 = \alpha(1-\alpha)(\mu_h d_l + \mu_l d_h)^2 + (1-\varepsilon)[\alpha\mu_h P_l - (1-\alpha)\mu_l P_h]^2 \tag{53}$$

Because $\varepsilon \leq 1$, there is still $(\partial\tilde{\lambda}/\partial W) > 0$

The second order derivative of $\tilde{\lambda}$ with respect to W is

$$\frac{\partial^2\tilde{\lambda}}{\partial W^2} = G^{\varepsilon}\varepsilon\alpha(1-\alpha)B^{\varepsilon-1} d_h^{\varepsilon\alpha-1} d_l^{\varepsilon(1-\alpha)-1} (\mu_h d_l + \mu_l d_h)^{1-\varepsilon} T_2^{-2} T_3 \frac{\partial\tilde{\lambda}}{\partial W} \tag{54}$$

where

$$\begin{aligned} T_3 \;\; = \;\; & \varepsilon T_2 d_l d_h[\alpha^2\mu_h\mu_l + (1-\alpha)^2\mu_h\mu_l + \alpha^2\mu_h^2(\frac{d_l}{d_h}) + (1-\alpha)^2\mu_l^2(\frac{d_h}{d_l}) \\ & - \mu_h\mu_l] + 2(1-\varepsilon)\mu_h\mu_l d_h d_l(\alpha\mu_h P_l - (1-\alpha)\mu_l P_h)^2 \end{aligned} \tag{55}$$

Consultation with Appendix A ought to reveal that $T_3 > 0$ because $(P_l + \tilde{\lambda}\mu_l)/(P_h + \tilde{\lambda}\mu_h) > \mu_l/\mu_h$, and $\alpha^2\mu_h^2 d + (1-\alpha)^2\mu_l^2(1/d) > 2\alpha(1-\alpha)\mu_h\mu_l$. Therefore, $\partial^2\tilde{\lambda}/\partial W \geq 0$. Hence when electricity market is imperfect, the marginal value of allowances is still convex in the underlying stochastic variable.

D Data Sources

I obtained annual data on initial allocations, holdings, transactions, and deductions (for emissions compliance purposes) of allowances from the EPA's allowance tracking system, spanning from 1995 to 2004 for all 263 Phase I affected coal-fired generating units. I merged this dataset with information collected from FERC Form 767. Form 767 is an annual survey of steam-electric unit operation and design, which covers the period 1996-2004. From Form 767 I took annual observations on unit nameplate capacity, the status of scrubber installation, electricity generation, load hours, fuel consumption and monthly fuel sulfur content and heat content. Using information about fuel consumption, fuel heat content and total generation, along with the emissions data derived from the EPA's allowance tracking system, I computed heat rate, heat input, and emission rate for each unit-year observation. Because data are missing for certain years, and because some units were not operating for an entire year, the number of observations varies from unit to unit.

The price information on fuel, electricity, natural gas, and allowances is obtained from different sources. To construct data on prices for low- and high-sulfur coal, the merged data set cited previously was then merged with FERC Form 423, which provides data on monthly cost and quality of coal for electric plants from 1995 to 2004. Form 423 records the physical quantity, Btu content, delivered cost, and sulfur content of each coal transaction at each electric plant. A SO_2 emission boundary of 2.5 pounds/MMBtu was used to distinguish low- and high-sulfur coal. This value was chosen so that the burning of low-sulfur coal meets Phase I standard on average. Fuel prices are calculated by dividing the delivered cost by the heat content of the fuel. Coal prices are missing for some plants when only one type of coal is purchased (so the price of the other coal is unobservable). To obtain the cost of coal that is not purchased by the plant, I use its price in the previous year as an approximation.

The electricity price data are drawn from the responses from electric utilities survey Form EIA-861, "the Annual Electric Power Industry Report." From this data set, I obtained the annual average industrial price for all years from 1995 to 2004 at the state level. Industrial prices are the most volatile and least protected by PUC regulation. The volatility of industrial prices is driven primarily by changes in fuel costs. However, because of the existence of long-term contacts, using industrial prices may underestimate actual price volatility in the spot market. Natural gas wellhead prices were collected from EIA historical database. SO_2 allowance price is calculated as the mean of two monthly price indices of SO_2 allowances that brokerage firms Cantor Fitzgerald and Fieldston report to the EPA.

Distance to fuel mines was provided by the EPA's Acid Rain Division. Data on regulatory status were collected from the Retail Wheeling & Restructuring Report, a state-by-state report of regulatory commissions, state legislation, and utilities activities related to retail competition, published quarterly by the Edison Electric Institute. These data are crosschecked with the LEAP Letter, published bimonthly by William A. Spratley & Associates, the National Regulatory Research Institute Web site, and the EIA's publication, "Status of State Electric Industry Restructuring Activity".

E Stochastic Dynamic Simulation Model Description

Here I present the computational details of the numerical simulation, programming language, hardware and software used. Consider the following optimization problem

$$V(P_{et}, B_t) \equiv \max_{l_t, h_t} \{P_{et} g(l_t, h_t) - c(l_t, h_t) + \beta E_t[V(P_{e(t+1)}, B_{t+1})]\} \tag{56}$$

$$s.t. \qquad \mu_l l_t + \mu_h h_t = A_t + B_t - B_{t+1} \tag{57}$$

$$0 \leq B_{t+1} \leq A_t + B_t \tag{58}$$

where $g_t = G(l_t^{\alpha} h_t^{1-\alpha})^{\varepsilon}$.

This is a multivariate optimization problem with two control variables (l, h). To simplify the problem, I derive the structural relationship between l, h and B based on first-order conditions. This derivation, which involves a calculation parallel to that given in Appendix C, leads to the following specifications:
(1) λ_t is fully characterized by

$$P_{et} = \frac{(A_t + B_t - B_{t+1})^{1-\varepsilon}}{G\varepsilon\alpha(1-\alpha)} d_{ht}^{1-\varepsilon\alpha} d_{lt}^{1-\varepsilon(1-\alpha)} (\mu_h d_{lt} + \mu_l d_{ht})^{\varepsilon-1} \tag{59}$$

(2) l_t and h_t are determined by the following two optimality equations

$$l_t = \frac{d_{ht}}{d_{ht}\mu_l + d_{lt}\mu_h}(A_t + B_t - B_{t+1}), \quad h_t = \frac{d_{lt}}{d_{ht}\mu_l + d_{lt}\mu_h}(A_t + B_t - B_{t+1}) \tag{60}$$

Noting that now the problem is simplified as we only need to search for the solution for B, and find optimal choices for l and h using Eqs. (59) and (60). Applying a grid search to obtain the initial guess for B ($B^* < 20$), I then specify a grid of 2001 points between 1 and 20 (0, 0.01, ..., 20) to compute the value function (56) at each time t and each state of P_e and B, beginning with the terminal period and working back to period 1. In this way, I compute the equilibrium time paths for B and λ, which jointly determine the optimal choices for l_t and h_t.

The program needed for the computation of the model was coded in C++ and complied to run on Windows-based machines. The whole simulation runs in one minute. All the code is available upon request from the author.

Figure 1-Annual Emission Cap, Aggregated Emissions and Banked Allowances: 1995-2004

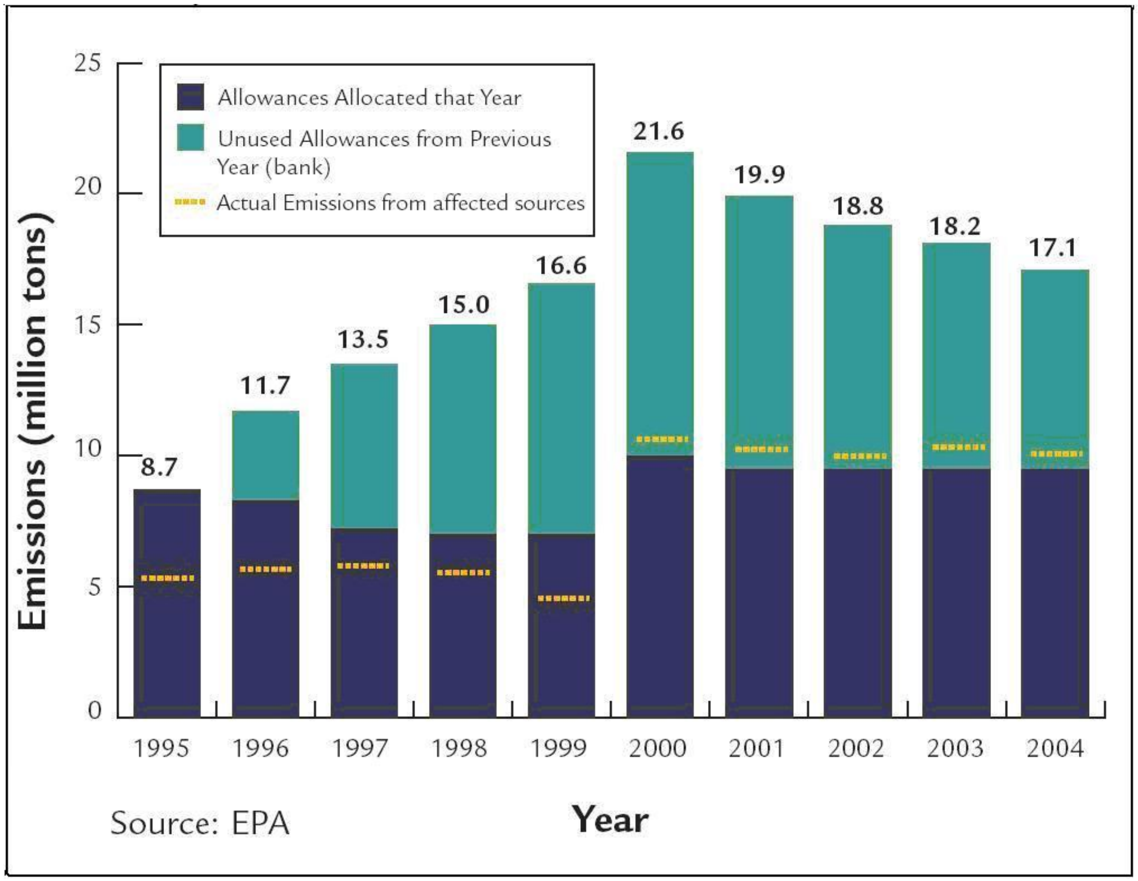


Figure 2-Electricity Market Volatility (A Simple Story)

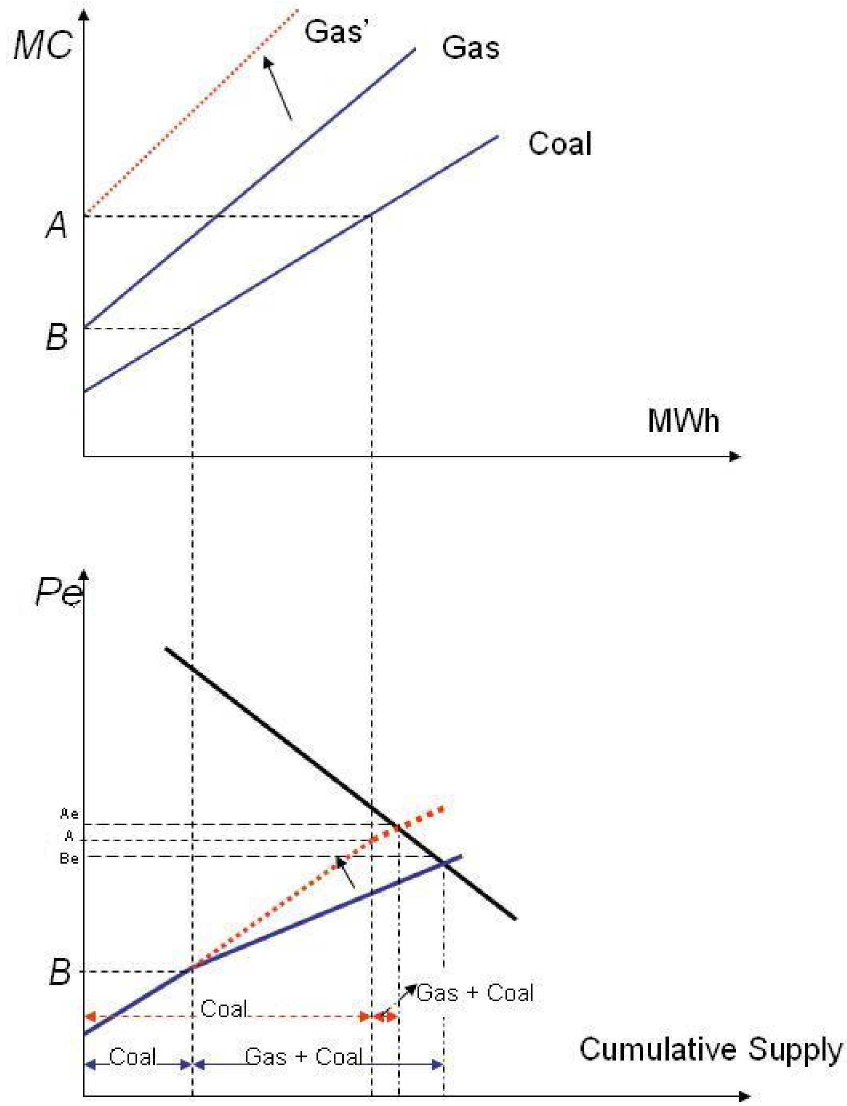
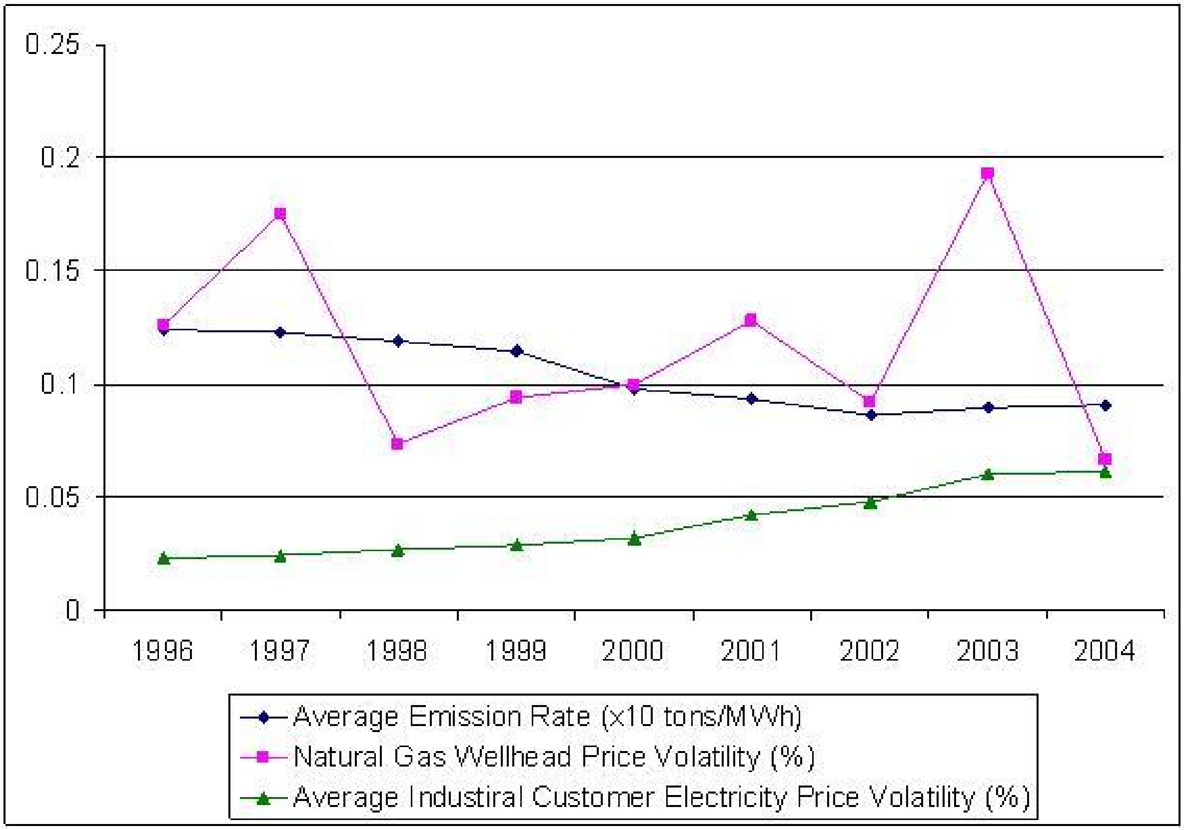
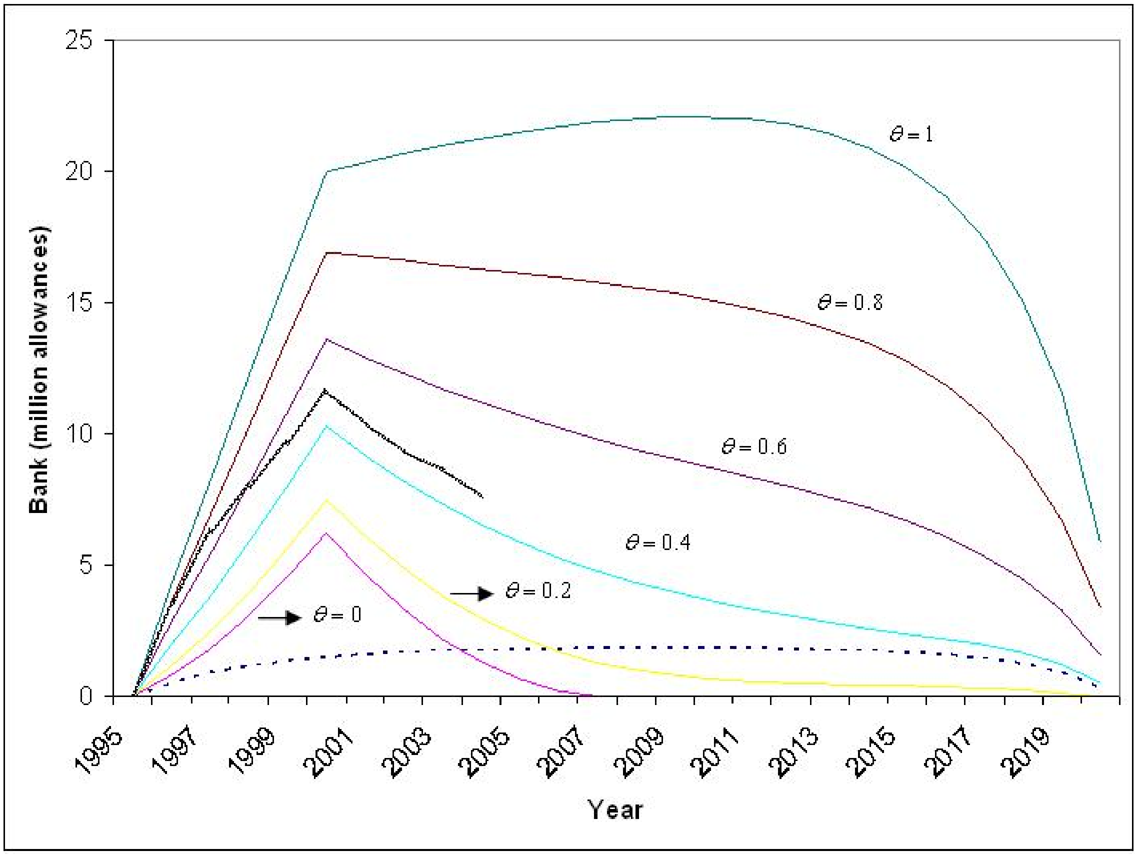


Figure 3-Trend in Emission Rates and Electricity Price Volatility: 1996-2004



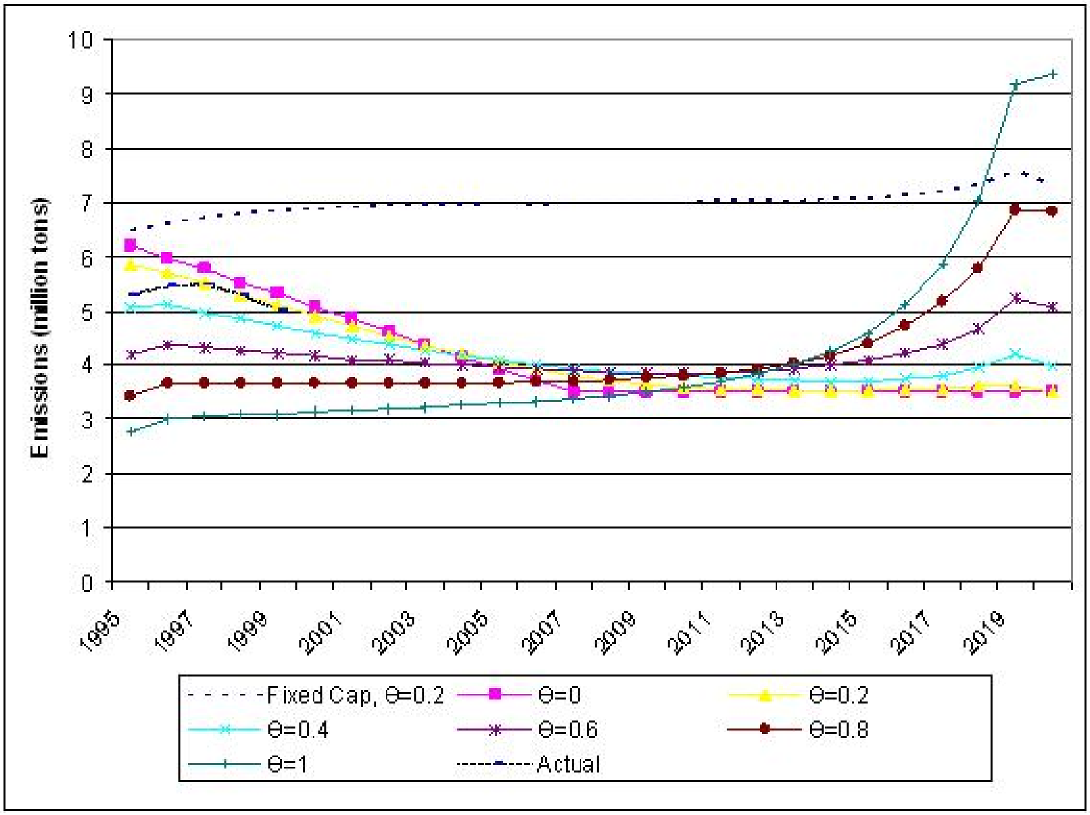
Note: Emission Rates are average annual emission rates of all Phase I affected units; natural gas and electricity price volatilities are calculated based on Eq.(31). Electricity price volatility is the average price volatility in the 21 states where Phase I units are located.

Figure 4-Total Amount of Allowances Banked under Different Price Volatilities



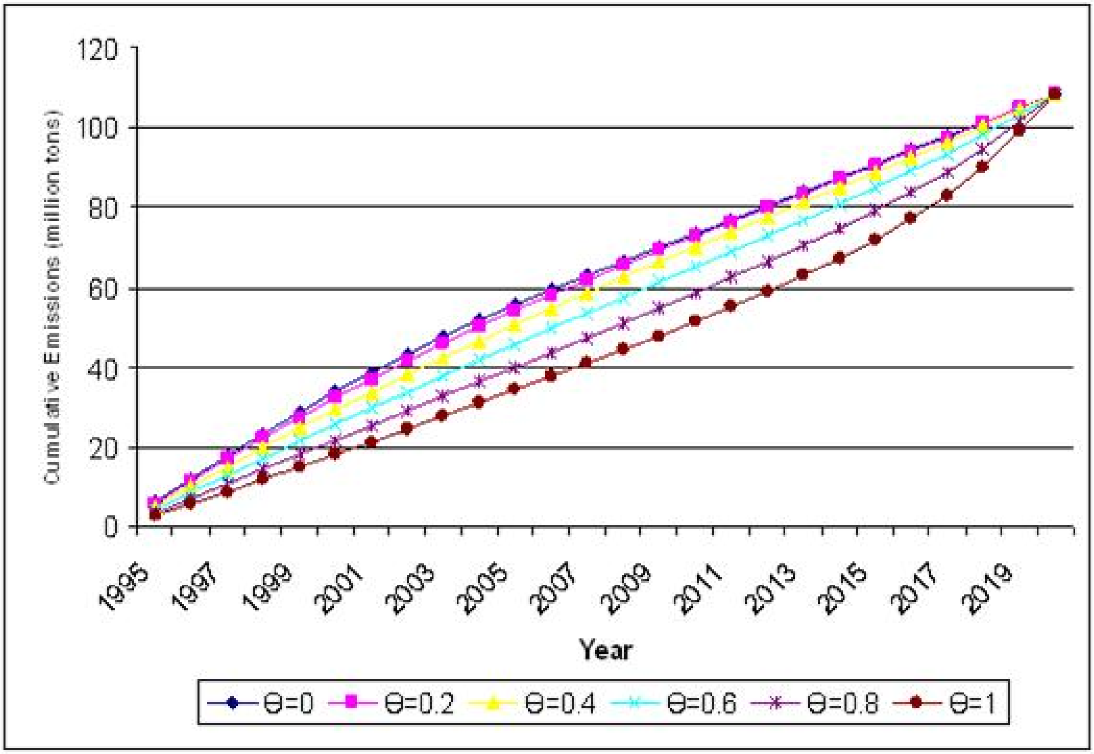
Note: The vertical axis describes the total amount of banked permits of the polluting industry during each year (million tons). θ is the mean-preserving spread of stochastic electricity prices. The dashed line corresponds to a scenario in which $\theta = 0.2$ and the emission cap remains constant at 7 million tons across all years. The shaded, fuzzy line tracks actual allowance stock in the SO_2 allowance market. The other lines correspond to a two-stage schedule of declining emission standards, with total emissions capped at 7 million tons from 1995 to 1999 and at 3.5 million tons during and after 2000.

Figure 5-Annual Emission Flow under Different Price Volatilities



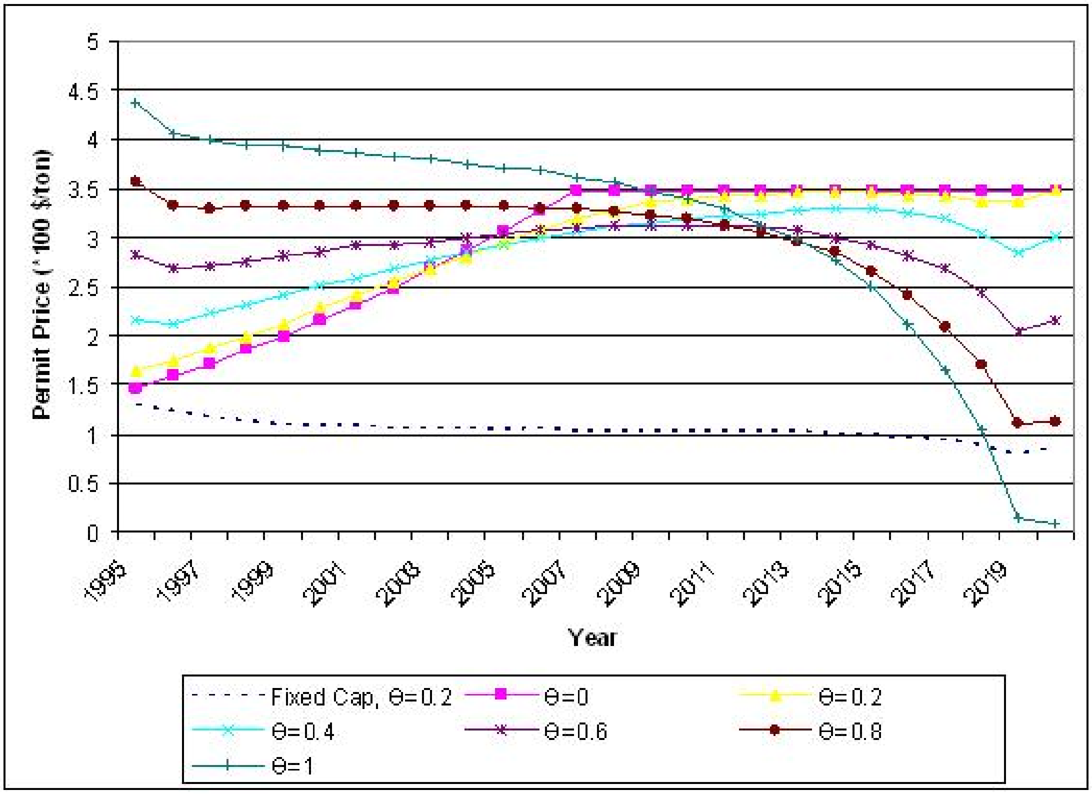
Note: The vertical axis describes the aggregate annual emissions of the industry (million tons). θ is the mean-preserving spread of the stochastic electricity price. The dashed line corresponds to a scenario in which $\theta = 0.2$ and the emission cap remains constant at 7 million tons across all years. The shaded, fuzzy line tracks actual allowance stock in the SO2 allowance market. The other lines correspond to a two-stage schedule of declining emission standards, with total emissions capped at 7 million tons from 1995 to 1999 and at 3.5 million tons during and after 2000.

Figure 6-Cumulative Emissions under Different Price Volatilities



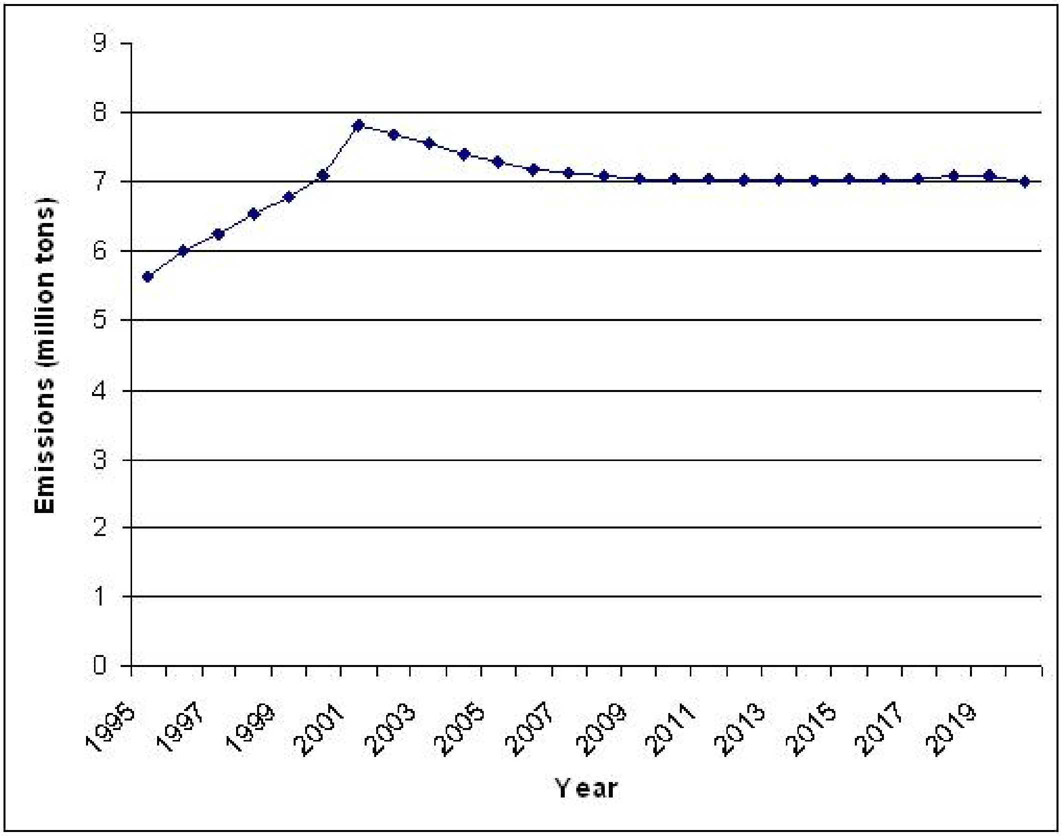
Note: The vertical axis describes the cumulative emissions during each year (million tons). θ is the mean-preserving spread of stochastic electricity price. The dashed line corresponds to a scenario in which $\theta = 0.2$ and the emission cap remains constant at 7 million tons across all years. The other lines correspond to a two-stage schedule of declining emission standards, with total emissions capped at 7 million tons from 1995 to 1999 and at 3.5 million tons during and after 2000.

Figure 7-Permit Price Path under Different Price Volatilities



Note: The vertical axis describes emission price (×100$/ton). θ is the mean-preserving spread of the stochastic electricity price. The dashed line corresponds to a scenario in which $\theta = 0.2$ and the emission cap remains constant at 7 million tons across all years. The other lines correspond to a two-stage schedule of declining emission standards, with total emissions capped at 7 million tons from 1995 to 1999 and at 3.5 million tons during and after 2000.

Figure 8-Emissions Path under Uncertainty Shock



Note: The graph describes the emission path under a scenario in which the price spread θ remained at 0.4 from 1995 to 2002, then dropped to 0.2 through 2020.

Table 1-Definitions of Symbols

l	low-sulfur coal
h	high-sulfur coal
g	electricity output
x	allowances bought ($x > 0$) or sold ($x < 0$)
P_e	electricity price
P_a	allowance price
P_l	low-sulfur coal price
P_h	high-sulfur coal price
A	allowances annually issued by the government
B	banked emission allowances
e	emission level
δ	coal's sulfur content
γ	conversion rate of sulfur to SO_2
μ	coal's SO_2 content
V	value of the firm
β	discount ratio
r	discount rate
ξ	ratio of marginal products of high- and low-sulfur coal
λ	shadow value of allowances
ϖ	convenience yield of banked allowances ($\varpi \geq 0$)
G	productivity factor of the electric industry
α	distribution weight of low-sulfur coal ($0 < \alpha < 1$)
θ	mean-preserving spread of stochastic electricity prices
ρ	elasticity of substitution parameter between low- and high-sulfur coal
φ	mark-up coefficient of a unit with market power in a electricity market ($\varphi \geq 1$)
W	exogenous stochastic demand shifter in the electricity market
ε	returns-to-scale parameter of a production technology $\varepsilon \leq 1$
s	allowed return on capital cost under ROR regulation
k	previously invested capital of an electric utility
π	a regulated firm's gross revenue net of operating expenses
ϕ	Lagrangian multiplier associated with ROR profit constraint. It reflects the extra profit a firm would get if the profit restriction is relaxed marginally. $0 < \phi < 1$ implies profit constraint is effective.
τ	extra return on allowances expenditures for a regulated utility compared to a competitive firm

Table 2-Description of Econometric Model Variables

Emissionrate	Total emissions (tons) divided by the total generation (MWh)
$\triangle P_e$	Annualized industrial electricity price volatility
VINTAGE	Boiler year on-line
AGE	Boiler age
AGE2	The square of AGE
HEATRATE	Annual heat input divided by the annual electricity output
WORKLOAD	Hours under load during a year
INITIAL	Annual allowances allocated by the EPA
CARRY	Allowances carried over to next period
MUNI	Municipally operated flag (1=yes, 0=no)
CAP	Generator 1989 nameplate capacity
CAP2	The square of CAP
RTE93	1993 boiler SO2 emission rate
DPRB	Unit's distance to the Powder River Basin
RETAILACESS	Plant being deregulated (1=yes, 0=no)
TRANSIT	Plant considered to be deregulated (1=yes, 0=no)
SCRUBBER	Boiler SO_2 scrubber flag (1=installed, 0=no)

Table 3-Summary Statistics

Variables	Obs.	Mean	Std.Dev.	Min	Max
Emissionrate(tons/MWh)	1595	0.011	0.007	0.00008	0.041
$\triangle$ Pe	1595	0.050	0.037	0.014	0.170
$\triangle$ Pe $_{1992}$	1595	0.033	0.025	0.015	0.108
$\triangle$ Pe $_{1990-1995}$	1595	0.044	0.025	0.019	0.119
Pa (dollars/ton)	1595	133	55	80	285
Pe (cents/KWh)	1595	4.21	0.95	2.68	9.54
Pl (cents/MMBtu)	1595	127.8	28	71.3	279
Ph (cents/MMBtu)	1595	126.7	43.2	76.7	418.6
Png (electric power price) (dollars/MMBtu)	1561	3.53	0.80	1.95	5.85
Png (wellhead price) (dollars/thousand cubic feet)	1595	2.97	1.34	1.77	5.45
AGE (years)	1595	36	8.1	18	55
HEATRATE(MMBtu/MWh)	1595	10.23	1.05	2.5	17.9
WORKLOAD (hours)	1595	7253.5	1077.8	792	8760
INITIAL (tons)	1595	17467	17877	144	192637
B (tons)	1595	11187	18472	0	155236
MUNI	1595	0.021	0.144	0	1
CAP (MW)	1595	356	254	75	1300
μ_l(lbs/MMBtu)	1452	1.64	0.68	0.41	2.98
μ_h(lbs/MMBtu)	1452	4.41	1.28	3	8.95
DPRB (miles)	1461	1063	327	87	1773

Table 4-Yearly Observations on Regulatory Status, Scrubber Installation and Fuel Switching/Blending

Year	Retail Access	Scrubber	Switch	No-blend
1996	28	23	34	21
1997	44	23	38	15
1998	58	23	35	23
1999	105	21	54	8
2000	87	21	56	15
2001	55	17	40	3
2002	55	19	65	11
2003	49	16	50	13
2004	49	17	36	7

Note: This table shows annual observations on the number of units affected by retail restructuring, those which installed scrubbers, switched to low-sulfur coal, or used only high-sulfur coal for production. Because many generating units were divestitured to non-utilities and were no longer reporting fuel purchasing costs to the public after 1998, the sample size shrinks over the years.

Table 5- Emission Rates Estimates (tons/MWh in Log Form)

Variables	(1)	(2)	(3)	(4)	(5)	(6)
△ Pe	-0.849**	-0.844**	-0.872**	-0.913**	-0.906**	-0.946**
	(0.372)	(0.372)	(0.418)	(0.332)	(0.383)	(0.370)
lnPa	-	-0.057	-	-0.634	-0.178*	0.020
		(0.078)		(0.614)	(0.102)	(0.077)
lnPe	-0.263	-0.232	-0.334	-0.310	-0.138	-0.492**
	(0.207)	(0.204)	(0.449)	(0.198)	(0.233)	(0.187)
lnPl	0.041	0.139*	0.017	0.147	0.034	
	(0.079)	(0.084)	(0.085)	(0.100)	(0.089)	
lnPh	0.107	0.047	0.125	-0.022	0.088	
	(0.082)	(0.084)	(0.081)	(0.079)	(0.089)	
RETAILACCESS	-0.049	-0.061	-0.050	-0.037	0.007	-0.019
	(0.047)	(0.047)	(0.047)	(0.045)	(0.059)	(0.050)
RETAILACESS△Pe	1.300**	1.380**	1.450**	0.635	-0.438	1.317**
	(0.483)	(0.489)	(0.541)	(0.413)	(0.526)	(0.482)
MUNI	-	0.176	- 0.088	0.319	-0.057	
		(0.334)	(0.275)	(0.332)	(0.298)	
MUNI△ Pe	2.175	2.108	2.133	1.017	2.055	1.897
	(1.433)	(1.434)	(1.417)	(0.828)	(1.461)	(1.439)
SCRUBBER	-2.447***	-2.144***	-2.316***	-1.995***	-2.183***	-1.828**
	(0.033)	(0.149)	(0.041)	(0.135)	(0.121)	(0.103)
AGE	-0.023	0.0005	-0.027	0.032	0.010	0.009
	(0.025)	(0.023)	(0.025)	(0.027)	(0.022)	(0.025)
AGE2	-0.0001	-0.0003	-0.0001	-0.0008	-0.0003	-0.0005*
	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
lnHEATRATE	0.120*	0.233**	0.109*	0.328**	0.241**	0.292**
	(0.071)	(0.093)	(0.068)	(0.108)	(0.093)	(0.098)
WORKLOAD	-0.056	-0.076	-0.065	-0.165	-0.023	-0.106
	(0.087)	(0.082)	(0.084)	(0.089)	(0.089)	(0.092)
INITIAL	1.19E-06	1.44E-06	1.23E-06	-8.60E-07	1.32E-06	1.52E-06
	(8.62E-07)	(8.74E-07)	(8.60E-07)	(1.70E-06)	(8.89E-07)	(8.33E-07)
lnCAP	-	-0.510	-	-0.616	-1.254	-0.724
		(0.909)	(0.890)	(0.966)	(0.943)	
lnCAP2	-	0.019	-	-0.029	0.096	0.039
		(0.077)	(0.075)	(0.078)	(0.084)	
DPRB						0.003*
						(0.001)
DPRB2						-3.06e-06*
						(1.66E-06)
DPRB3						9.29e-10*
						(5.30E-10)
LSPremium						0.090***
						(0.028)
Constant	-4.133***	-2.808	-	-	-0.803	-1.975
	(0.755)	(2.870)		(3.189)	(2.686)	
R^2	0.325	0.672	0.265	0.683	0.724	0.658
Observations	1595	1595	1560	1243	1300	1457
N	208	208	207	208	146	192

Note: The dependent variable is ln(emissionrate) (tons/MWh in log form). Columns (1) and (2) report results from estimating equation (23) via fixed and random effects models. A Hausman test does not reject the null hypothesis that there is no systematic difference between fixed and random effects estimations. The test statistics are $\chi^2(19) = 3.66$, P-value = 0.9999. Column (3) reports IV/GMM estimation using natural gas electric power and wellhead prices as instruments for electricity and SO_2 prices. The Cragg-Donald F statistic of weak identification is 56.707. The Hansen J test statistic of overidentification of all instrument is 5.711, the P-value is 0.0169. The Anderson cononical correlation LR test statistic of underidentification is $\chi^2 = 110.48$, and the P-value is 0.0000. Column (4) reports random effects estimation based on data from 1997 to 2003. Column (5) reports GLS random effects estimation based on a balanced panel containing 146 units that remain in the sample from 1996 to 2004. Column (6) reports estimation results from equation (23) using units' distance to PRB coal and low-sulfur coal premium as proxies for low- and high-sulfur coal prices. Standard errors clustered by unit are reported in parentheses. *** indicates significant at the 1% level; **indicates significant at the 5% level; * indicates significant at the 10% level. Reported R^2 is the adjusted R^2 for fixed and random effects models, and centered R^2 for IV/GMM model. N is the number of clusters.