Implementing a Time- and Location-Differentiated Cap-and-Trade Program: Flexible Nitrogen Oxide Abatement from Power Plants in the Eastern United States

by

Katherine C. Martin

B.A., Physics
Reed College, 2000

S.M., Technology and Policy
Massachusetts Institute of Technology, 2006

SUBMITTED TO THE ENGINEERING SYSTEMS DIVISION IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN TECHNOLOGY, MANAGEMENT, AND POLICY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

Signature of Author ....
Engineering Systems Division
July 2007

Certified by ....
David H. Marks
Goulder Family Professor of Civil and Environmental Engineering and Engineering Systems
Thesis Committee Chair

Certified by ....
A. Denny Ellerman
Senior Lecturer, Sloan School of Management

Certified by ....
Paul L. Joskow
Elizabeth and James Killian Professor of Economics

Certified by ....
Kenneth A. Oye
Associate Professor of Political Science and Engineering Systems

Accepted by ....
Richard de Neufville
Professor of Engineering Systems
Chair, Engineering Systems Division Education Committee

Implementing a Time- and Location-Differentiated Cap-and-Trade Program: Flexible Nitrogen Oxide Abatement from Power Plants in the Eastern United States

by
Katherine C. Martin

Submitted to the Engineering Systems Division on July 13, 2007
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy in Technology, Management, and Policy

ABSTRACT

Studies suggest that timing and location of emissions can change the amount of ozone formed from a given amount of nitrogen oxide (NO_X) by a factor of five (Mauzerall *et al.* 2005). Yet existing NO_X cap-and-trade programs require stationary sources in the Eastern U.S. to reduce emissions without reference to timing or location. This work is part of a larger study on whether a NO_X cap-and-trade program that differentiates across emissions by time and location could reduce ozone concentrations more cost-effectively than simple aggregate reductions in the NO_X cap in the Eastern United States.

To gauge possible gains relative to existing regulations, this work examines compliance data from coal power plants in 2002 and 2005 to estimate the effectiveness of existing un-differentiated regulations. It finds that some plant operators chose to remain under aggregated caps by emitting less NO_X during early summer months when effects on ozone formation are low and emitting more NO_X during late summer months when effects on ozone formation are great. This behavior was at once individually rational, environmentally damaging, and perfectly legal.

To evaluate potential challenges to implementation, the study assesses the technical feasibility and the distributional effects of spatially and temporally differentiated regulatory systems.

* Are power plants in the Eastern U.S. technically capable of reducing NO_X emissions in response to incentives that changed in time and by location given network constraints? To address these questions, this work used a zonal model based on an abstract network graph and optimal power flow simulations to estimate potential short-term NO_X reductions and associated costs from redispatch of power plants in the original Pennsylvania-New Jersey-Maryland (PJM) power system. Both methods estimated that power plants could respond with hourly NO_X reductions of between 15 and 30% and that network constraints had little effect.

* Are the distributional effects of a differentiated regulation likely to motivate and/or enable legal challenges that could undercut such a program? The distributional effects of differentiated regulation would depend on the timing and locations of reductions, and legal challenges could constrain implementation. But the inability of un-differentiated regulations to fully solve ozone problems, combined with scientific and economic justifications, and the ability of power plants to respond, justify further inquiry into the feasibility of differentiation.

Thesis Committee:
A. Denny Ellerman (Research Supervisor), Senior Lecturer, Sloan School of Management
Paul L. Joskow (Research Supervisor), Elizabeth and James Killian Professor of Economics
David H. Marks (Chair), Goulder Family Professor of Civil and Environmental Engineering and Engineering Systems
Kenneth Oye, Associate Professor of Political Science and Engineering Systems

Table of Contents

Chapter 1 – Introduction

Ground-level ozone is a pollutant that damages public health and the environment. To protect public health and welfare, Congress mandated that the Environmental Protection Agency (EPA) set a National Ambient Air Quality Standard (NAAQS) for ozone in the United States in the Clean Air Act of 1970. The EPA implemented the ozone standard in 1971 and made it more stringent in 1997.[1] Some areas of the United States have had difficulty achieving the ozone standard despite regulations requiring considerable reductions in the precursor emissions that contribute to ozone formation – nitrogen oxides (NO_X) and volatile organic compounds (VOCs). Areas of the Eastern U.S. have found it particularly difficult to achieve the ozone air quality standard even though seasonal NO_X cap-and-trade programs for stationary sources have reduced NO_X emissions substantially since 1999.[2] Recent regulation requires decreases in the seasonal cap on NO_X emissions from stationary sources in 2010 and 2015, but the EPA still does not expect all areas of the Eastern U.S. to attain the ozone standard by 2015 (U.S. EPA 2006b).[3] One potential reason for the persistent nonattainment is that ozone formation depends not only on the quantities of precursor emissions but on their timing and location.

1.1. Problem statement and contribution

This dissertation contributes to a broader research effort that is seeking to determine whether it would be more cost-effective to directly address the temporal and

---

[1] The EPA is currently considering a further increase in stringency of the ozone air quality standard.

[2] The Ozone Transport Commission (OTC) NO_X Budget Trading Program (implemented in 1999) and State Implementation Plan (SIP) NO_X Budget Trading Program (implemented in 2003) have reduced NO_X emissions from stationary sources in the Eastern U.S. by 72% from 1990 levels (U.S. EPA 2006b).

[3] The Clean Air Interstate Rule (CAIR) se 70 *Federal Register* 25162 (May 12, 2005).

locational impacts of NO_X emissions from power plants on ozone formation in the Eastern United States rather than to further reduce the aggregate cap of the seasonal cap-and-trade program now in place.[4] To contribute to the research addressing this question, this dissertation first examines scientific and economic reasons to use a temporally and spatially differentiated cap-and-trade program to regulate NO_X emissions from power plants. Then, using historical data to test the hypothesis that the operators of some power plants reduced emissions more during times when ozone formation was least likely under un-differentiated cap-and-trade programs, this dissertation provides a detailed example of a potential benefit of differentiation.

While there are scientific and economic justifications for the use of differentiated regulations, there are implementation challenges as well. These challenges are the reason that differentiation in regulations and policies has not been implemented to address the ozone problem. By simulating one of the electric power systems in the Eastern United States, this dissertation also analyzes the feasibility of overcoming two implementation challenges. The first is whether power plants could respond to a differentiated regulation given network constraints – the finding is that they could. The second is whether the distributional effects of a differentiated regulation could constrain its implementation or reduce its effectiveness. This constraint on implementation may have a slightly larger effect in the case of a differentiated regulation but would not necessarily negate the potential benefits of differentiation.

An important finding of this work was that network constraints would not greatly constrain the potential NO_X reductions from power plants. A misperception prior to this study was that network constraints would limit the redispatch[5] of power plants as a

---

[4] The research is a project of the Center for Energy and Environmental Policy Research at MIT. See also: Martin, K.C., P.L. Joskow, and A.D. Ellerman (2007), "Time and Location Differentiated NO_X Control in Competitive Electricity Markets Using Cap-and-Trade Mechanisms," *Center for Energy and Environmental Policy Research Working Paper.*

[5] "Redispatching" power plants implies changing which particular generating units fill electricity demand at a given time. Power plants are typically economically dispatched, which means that the lowest cost generating units are used

potential response to a differentiated cap-and-trade program. The overall conclusion of this work is that the evidence of differentiation's potential benefits and the fact that power plants could respond with NO_X reductions to differentiated incentives justify further inquiry into the feasibility of using a differentiated cap-and-trade program to address the ozone problem in the Eastern United States. The temporal and spatial features of the ozone problem are particularly important. But other environmental problems, like particulate matter pollution, have some similar characteristics. If a differentiated regulation could succeed for ozone pollution, it might also be applied in other cases.

1.2. Summary

The atmospheric chemistry literature shows that NO_X emissions reductions at some times and locations affect ozone formation more than at others. The contribution of NO_X emissions to ozone formation depends on sunlight, temperature, and concentrations of VOC emissions. All of these factors vary in time and by location. Sunlight and temperature vary with the weather. Biogenic sources like oak trees emit about 60% of the VOCs in the Eastern United States. The VOCs emissions from these sources not only vary by location but also in time because many natural sources emit more VOCs when it is hot and sunny. NO_X emissions tend to cause more ozone during hot, sunny conditions in areas with high concentrations of VOC emissions. For example, one study found that NO_X emitted under these conditions caused five times more ozone to form as the same amount of NO_X emissions under conditions less conducive to ozone formation (Mauzerall *et. al.* 2005). Winds can also transport NO_X emissions and cause ozone formation displaced in time and space from the original source. This further complicates the relationships between emission sources and ozone concentrations. This literature suggests that the effectiveness of regulations like the NO_X cap-and-trade programs in the

---

first to fill demand provided that network constraints and other system requirements are met. These simulations assume that a price on NO_X emissions changes the relative costs of generating units due to differences in the units' NO_X emission rates and therefore a high NO_X price would change the order in which the units are dispatched.

Eastern United states could be limited by their failure to specifically address the temporal and spatial aspects of the contribution of NO_X emissions to ozone formation.

The environmental economics literature discusses the potential efficiencies associated with using differentiated regulations to address environmental problems with time and spatial dimensions, like ozone. A differentiated regulation could improve cost-effectiveness compared to a un-differentiated regulation by providing incentives for emission sources to reduce NO_X when they would most likely mitigate ozone formation and by not requiring emission reductions in times and locations when they would not improve air quality. In the particular case of ozone pollution in the Eastern United States, this dissertation presents an analysis of historical data from coal power plants to illustrate this potential benefit of using a differentiated cap-and-trade program. Under the un-differentiated cap-and-trade program now in place, the operators of some power plants reduced NO_X emissions more during times when ozone formation was least likely. This behavior was consistent with economic incentives provided by the NO_X cap-and-trade program and the wholesale electricity markets in which the power plants participated. An effective differentiated regulation could provide the strongest incentives for NO_X reductions when they would be most likely to mitigate ozone formation. Although the specific emission reductions required by a differentiated cap-and-trade program could be more costly than those undertaken in the un-differentiated case, the differentiated regulation would improve cost-effectiveness by not requiring ineffective reductions.

Compelling scientific and economic arguments support the use of a differentiated regulation rather than further reductions in the cap of the current regulation, but there are challenges to implementation. This dissertation examines two of these challenges. The first is technical and economic: could power plants respond to incentives for short-term NO_X reductions that changed in time and by location and what would this type of flexible abatement cost? The second challenge is political, economic, and legal: the distributional effects of a differentiated program could constrain its effectiveness. Small changes in the definitions of the locations and timing of required emission reductions could have large benefits for industries or states. These affected parties could use political influence or legal disputes over the modeling necessary to support a differentiated regulation to shape the details of the regulation and this could limit the regulation's effectiveness.

This dissertation analyses the first technical and economic challenge in detail and proposes that it is tractable. Simulations of potential NO_X reductions from the redispatch of power plants suggest that it would be technically feasible to reduce emissions nontrivially in response to incentives that changed in time and by location. The initial misperception that motivated this work was that network constraints would prevent the redispatch of power plants as a short-term response to differentiated incentives for NO_X reductions. The simulations suggested that network constraints would not greatly constrain the potential NO_X reductions because the exchange of generation that reduced NO_X occurred locally and did not require large-scale transfers of power across the network. The estimated costs of these NO_X reductions were comparable to other more conventional approaches (like requiring NO_X control technologies on some uncontrolled power plants).

The second implementation challenge is more difficult to resolve. The simulations of potential NO_X reductions from the redispatch of power plants illustrated that the distributional effects of a differentiated regulation would depend on the specifics of when and where emission reductions were required. The distributional effects of environmental regulations often motivate industry and state governments to contest or try to influence the specific details of the regulations because small changes in a regulation can yield large benefits. They typically use claims of fairness and disputes over the uncertainties in the modeling used to support a regulation as the legal means to influence the details of a regulation. In these ways, a differentiated regulation would not differ greatly from other environmental regulations. But two factors could amplify the effects of legal disputes on differentiated regulation compared to other environmental regulations. First, if the uncertainties associated with the modeling required to support a differentiated regulation were greater because of the need for more detailed spatial resolution, for example, interested parties could more easily contest the regulation by using their own models to show that their emissions do not affect the air quality in targeted areas. Second, if small changes in the locations, times, or quantities of the required emissions reductions could hinder a differentiated regulation's effectiveness then affected parties' attempts to shape the regulation's details to their benefit could be a problem.

While recognizing that political and legal processes could constrain the effective implementation of a differentiated cap-and-trade program, this dissertation focuses on the questions of whether or not such a regulation could be beneficial in the case of ozone pollution in the Eastern United States and on whether it would be technically feasible from the perspective of obtaining flexible emission reductions from power plants. The primary reason for this focus is that the conception of an implementable differentiated regulation is still its early stages. The environmental economics literature has recommended that regulations differentiate between emissions by providing the strongest incentives to reduce the most environmentally damaging (or problematic) emissions since the early 1970s (Montgomery 1972 and Mendelsohn 1986). But the complexities of implementing differentiated cap-and-trade programs and related high transaction costs have impeded their use (e.g. Tietenberg 1995). One such complexity in the case of ozone is that the contribution of NO_X emissions to ozone formation fluctuates with the weather, which means that weather and atmospheric chemistry forecasting must be used to alter incentives for NO_X reductions. The atmospheric chemistry literature suggests that advances in air quality forecasting and modeling may now mitigate this challenge and further work is currently being undertaken to validate these results.[6]

This dissertation assesses another specific technical challenge. An impediment specific to applying a differentiated regulation to NO_X emissions from power plants is the perception that thermal and security constraints in the electric power network would limit potential short-term reductions in NO_X emissions especially during times of peak electricity demand. The results of simulations that used two different methods suggest that network constraints do not greatly limit the potential for short-term emission reductions from power plants. Simulations of the redispatch of power plants in the portion of the PJM electric power network that covers Pennsylvania, New Jersey, Maryland, Delaware and the District of Columbia ("Classic PJM") suggest that it is technically and economically feasible to reduce NO_X emissions by between 6 tons (or 15%) in the highest demand hours and 8 tons (or 30%) in average demand hours. This dissertation advances the discussion of whether ozone regulations could make use of scientific information about the importance of the timing and location of NO_X emissions, as well as advances in air quality forecasting to improve the cost-effectiveness of the regulation of NO_X emissions from power plants: power plants could respond to a differentiated regulation with short-term NO_X reductions.

Future work will show whether or not the potential emission reductions from power plants would be enough to help areas in the Eastern United State attain the ozone air quality standards. If these reductions could help improve air quality, it would be worthwhile to consider other implementation challenges, like political constraints, in more detail. If, however, the potential short-term reductions from power plants are not sufficient to drastically improve air quality the focus of analysis could move away from

---

[6] This work is currently being undertaken at the Center for Energy and Environmental Policy Research at MIT.

power plants to assess the contribution of other sources of NO_X emissions like mobile sources to ozone air quality problems.

1.3. Overview of chapters

This dissertation is organized as follows. Chapter 2 reviews the atmospheric chemistry literature, which suggests that nonlinearities in the chemistry of ozone formation create the need for a temporally and spatially differentiated regulation for NO_X emissions. The chapter also discusses the feasibility of forecasting ozone concentrations and categorizing areas or times important for NO_X reductions. Both of these factors could help enable the implementation of a differentiated regulation. In order to motivate the need for reductions in ozone, Chapter 2 also summarizes the environmental and health effects of ozone and its precursor emissions. This discussion explains the justifications for policies that limit ozone concentrations as well as the reasons other than ozone to limit NO_X and VOC emissions. The reasons to reduce NO_X emissions involve tradeoffs. Section 2.2.2 discusses an example in which reductions in NO_X emissions reduced ozone concentrations but increased particulate matter concentrations. Tradeoffs like this are a challenge for environmental policy in the U.S. because of the statutory framework for air quality. Section 2.4.3 discusses whether a differentiated regulation could help regulators address tradeoffs while working within the statutory framework. This discussion also explains why the statutory framework for air quality regulation in the United States requires a focus on achieving the air quality standards for ozone rather than on balancing the costs and benefits of reducing ozone in each area of the country.

Chapter 2 also examines the sources of ozone precursor emissions and the historical trends in those emissions. One barrier to implementing a differentiated regulation for sources other than large stationary sources is that few data are available on how emissions from other sources (like mobile sources) vary in time and by location. The joint discussion of the science and policy backgrounds of the ozone problem in Chapter 2 helps explain why policies to reduce ozone concentrations have not been entirely successful even though they have caused dramatic reductions in ozone precursor emissions.

Chapter 3 begins by reviewing the environmental economics literature that discusses differentiated regulations. In doing so it explains the economic rational for using a differentiated approach to regulate NO_X emissions. The environmental economics literature also explains why differentiated regulations are difficult to implement. A review of this literature helps clarify the reasons why differentiated regulations have not been used to address the ozone air quality problem despite persistence of the problem and the scientific and economic justifications for differentiation.

Chapter 3 then discusses the NO_X cap-and-trade programs that have regulated emissions from stationary sources in the Eastern United States. It reviews studies that have analyzed the effectiveness of these regulations. The studies find that the regulations have been effective in achieving their targeted NO_X reductions and in doing so at low cost. However, although ozone air quality in the Eastern U.S. has improved substantially since the implementation of these regulations, the problem has not been completely solved. In fact, some studies suggest that the costs of the NO_X reductions achieved by these programs cannot be justified by improvements in ozone air quality alone. However, the programs' costs can be justified by the combination of reductions in particulate matter and ozone pollution. This explains one of the arguments against a differentiated regulation for NO_X emissions raised by environmental groups: reductions in NO_X emissions have other benefits and are therefore always justified. The tradeoffs between reducing ozone and particulate matter pollution discussed in Section 2.2.2 suggest that this argument is too simple. Also, NO_X reductions certainly have benefits beyond reducing ozone concentrations, but federal law mandates that all areas achieve the ozone air quality standards. The EPA's air quality modeling shows that the reductions in NO_X

emissions required by the Clean Air Interstate Rule, which focuses on reducing particulate matter pollution as well as ozone, will not guarantee that all areas of the Eastern U.S. attain the ozone air quality standards by 2015 (U.S. EPA 200b). This means that despite the other benefits of NO_X reductions, a careful approach that considers ozone formation chemistry is needed if the ozone air quality standards are to be attained.

Chapter 4 turns to the NO_X emission characteristics of coal power plants in order to examine the hypothesis that the operators of coal power plants have counterproductive incentives to reduce NO_X emissions at the times when ozone formation is least likely under the summertime cap-and-trade programs in the Eastern United States. A simple economic model that builds from the technical NO_X emission characteristics of coal power plans describes why some plant operators might have incentives to reduce NO_X emissions more when power prices are lower during the cooler parts of the summer (e.g. May and June compared to July and August). This is potentially problematic because peak ozone concentrations are less likely during the cooler parts of the summer. Historical data on emissions, generation and electricity, fuel, and NO_X prices were then used to estimate counterfactual emissions for what the emissions of coal power plants would have been without the cap-and-trade programs. The operators of some power plants reduced emissions more during May and June, which is consistent with the incentives. However, at the aggregate level nearly equal reductions occurred during the entire summer. This analysis illustrates a potential benefit of a differentiated regulation: it could help ensure that all sources undertake the most effective NO_X reductions.

Chapter 5 presents the results of the simulations that used two methods to estimate the potential short-term NO_X reductions from power plants under a differentiated cap-and-trade program. The simulations were performed for the electric power network that covers Pennsylvania, New Jersey, Maryland, Delaware, and the District of Columbia (called "Classic PJM"). Both simulation methods – a zonal model based on an abstract representation of the network and Optimal Power Flow (OPF) simulations – accounted for thermal and security constraints. Both methods suggested that power plant redispatch could reduce hourly NO_X emissions from power plants by between 6 tons (or 15%) on the highest demand days and 8 tons (or 30%) on average demand days. Both methods also suggested that network constraints did not constrain the potential reductions largely because the exchanges of generation that caused the reductions occurred locally. (Chapter 5 also compares the different methods in more detail.) Other Eastern U.S. power systems share key characteristics with Classic PJM that suggest these results may hold more generally.

Chapter 5 also discusses the results of simulations that estimated the relationship between assumed NO_X prices and abatement from redispatch. The costs are high but comparable to other strategies for NO_X reductions being considered by state governments. The same simulations suggested the distributional effects of a differentiated cap-and-trade program applied to power plants would not only depend on the allocation of NO_X emission allowances, as is typically the case with cap-and-trade programs, but also on the timing and locations of the required NO_X reductions.

Chapter 6 summarizes and offers opportunities for future work.

Chapter 2 – The Science of Ground-level Ozone Pollution and the History of U.S. Ozone Policy

This chapter first presents background information on the science of ozone formation in order to explain why the atmospheric chemistry literature suggests that a differentiated approach to controlling NO_X emissions is needed to reduce ozone concentrations (Section 2.1). This literature also suggests it may be possible to forecast and categorize ozone episodes, which could aid implementation of differentiation. Section 2.2 then discusses the damages caused by ozone and ozone precursor emissions, which justify policies that limit ozone and its precursor emissions. Section 2.3 discusses the sources of and historical trends in these pollutants. This discussion explains that regulators have focused on reducing NO_X emissions in the Eastern U.S. because of high biogenic concentrations of VOC emissions. It also explains why a cap-and-trade program for stationary sources might be easier to implement than one for mobile sources: few data are available on how NO_X emissions from mobile sources vary in time and by location.

Section 2.4 discusses the federal ozone air quality standards and the areas of the country that have had difficulties attaining them. This section also discusses the statutory framework for air quality policy and why this framework requires a focus on attaining air quality standards, rather than on the costs and benefits of reducing ozone concentrations. The joint discussion of the science and policy backgrounds of the ozone problem helps explain why policies to reduce ozone concentrations have not been entirely successful even though they have caused dramatic reductions in ozone precursor emissions. The history of NO_X regulations in Section 2.5 shows that state and federal regulators acknowledged the chemistry of ozone formation when they switched their focus away from VOC emissions to NO_X reductions in the Eastern United States. But it also shows that practical considerations have led policy makers to implement straightforward regulations that have not fully accounted for the impacts of the time and location of NO_X emissions on the formation of ozone in areas of concern.[7]

2.1. Basic ozone chemistry

Ground-level ozone forms in the lowest level of the earth's atmosphere, the troposphere.[8] The basic reactions that form ozone are reactions of VOCs that create compounds that react with nitric oxide (NO), the latter being emitted from the burning of fossil fuel, to form nitrogen dioxide (NO_2) (see Borrell 2003 for more details). The NO_2 created by these reactions, absorbs sunlight during the daytime. This creates an extra oxygen atom that can combine with O_2 to form ozone (O_3):

$$NO_2 + \lambda\,(400\text{ nm}) \rightarrow NO + O$$

$$O + O_2 \rightarrow O_3$$

In areas of high concentrations of NO_X, the concentrations of ozone are kept low by a reaction called the titration reaction:

$$NO + O_3 \rightarrow NO_2 + O_2$$

These three reactions and the resulting ozone concentrations depend on the relative concentrations of VOCs and NO_X ($NO + NO_2$), on the temperature, and on the amount of sunlight. Areas, or times, characterized by high concentrations in NO_X and relatively low concentrations of VOCs, are said to be VOC-limited. This means that a reduction in VOCs will likely reduce ozone-formation but a reduction in NO_X will stop the titration reaction and actually *increase* ozone concentrations. Most non-urban areas in

---

[7] Areas of concern include counties that are out of compliance with the federal air quality standard for ozone as well as highly populated areas with, even occasional, high ozone concentrations.

[8] See the EPA's "Basic Concepts in Environmental Sciences," Chapter 6: Ozone at http://www.epa.gov/eogapti1/module6/ozone/formation/formation.htm.

the Northeastern U.S. are NO_X-limited, meaning that reductions in NO_X will decrease ozone formation, although this does vary with time, as the amount of sunlight, wind, and the temperature vary. Additionally, in the Eastern U.S., ozone's lifetime is typically less than, or up to, two days (Fiore *et. al.* 2002).[9] This is long enough to make the transport of ozone and its precursors to downwind areas a problem. The wind in the Eastern U.S. typically transports the pollutants from west to east.

The right combinations of precursor emission levels, sunlight, and wind occasionally produce periods of high ozone concentrations (above 0.08 ppm), called ozone episodes, which typically last for a few hours up to a few days – and ozone concentrations typically dip below the standard during the night. As the public health impacts from ozone exposure worsen with increasing concentration, it is these episodes that are of particular importance in air quality policy.

2.1.1. Policy implications of ozone chemistry

Experience and the literature have highlighted the policy implications of the chemistry of ozone formation. For example, the counterintuitive relationship that very high concentrations of NO_X can suppress ozone formation explains the "weekend effect" in the Los Angeles air basin: ozone concentrations were higher on weekends when NO_X emissions from diesel trucks were lower (CARB 2004). Also, Ryerson *et. al.* (2001) found that ozone is less likely to form in the concentrated plumes from the largest power plants compared to the plumes from smaller plants. In addition, reductions of NO_X from power plants located near natural sources of VOCs, like oak forests, reduce ozone formation more than reductions from those far from VOC sources (Ryerson *et. al.* 2001, Chameides *et. al.*1988). Ryerson *et al.* (2001) summarize that a reduction of one ton of NO_X from a dilute power plant plume into an area with high ambient VOCs

---

[9] The lifetime of ozone in the troposphere in general is longer, 6 to 10 days, but factors like a shallow mixing layer and interactions with biogenic VOCs make ozone's lifetime shorter over the Eastern United States (Fiore *et al.* 2002).

concentrations would result in at least twice the amount of reduction in ozone formation compared to a reduction of one ton of NO_X from a concentrated plume in an area with low ambient VOC concentrations. These authors suggest that these relationships should be taken into account when designing NO_X permit trading programs in the Eastern United States.

More recent papers have used techniques that integrate atmospheric chemistry modeling with economic and demographic data in order to link the variable role of NO_X emissions in ozone formation to human exposure and health impacts. Mauzerall *et. al.* (2005) examined differences in health effects of ozone formation and exposure from NO_X emissions from large point sources at different locations and times. They chose times and locations that captured relevant ranges of variation in temperature and local biogenic VOC emissions. They found that the ozone produced from the same amount of NO_X emissions at these different times and places can vary by up to a factor of five. The public health impacts of the NO_X also depend on locational variations in demographics that influence exposure (Mauzerall *et. al.* 2005). Tong *et. al.* (2006) used similar techniques to study the ozone-caused NO_X damages around Atlanta. They found that the marginal damages of NO_X emissions vary greatly across the Atlanta metropolitan area because of ozone formation chemistry, including the effects of the titration reaction.

2.1.2. Predicting and categorizing the effects of NO_X on ozone episodes

A time- and location-differentiated cap-and-trade program could incorporate the role of meteorology and chemistry in the conversion of NO_X to ozone if two conditions concerning predicting and categorizing ozone episodes hold. The first is that weather and atmospheric chemistry forecasting can predict the conditions conducive to ozone formation with sufficient accuracy and lead-time (at least 48 hours) to influence electricity markets (or other decisions, such as where the regulations applied to vehicles or other sources). The second condition is that the spatial zones and time intervals in which the surrender ratio for the NO_X emissions permits would be varied can be identified with sufficient regularity that a reasonably simple and stable system of differentiated permit exchange rates triggered by transparent weather and atmospheric chemistry indicators can be implemented.

The literature suggests that these conditions are feasible.[10] For example, slow-moving, high-pressure systems drive the worst ozone episodes in the Eastern United States (NRC 1991 citing RTI 1975, Decker *et al.* 1976). This means that forecasting ozone episodes requires forecasting these high-pressure systems. The latter are generally predictable with a lead-time of 3 to 5 days (NRC 1991 citing Chen 1989, van den Dool and Saha 1990). The literature suggests that categorizing the relationships between NO_X emissions, meteorology, and ozone in defined geographic areas is possible. For example, Lehman *et. al.* (2004) studied rural and suburban ozone concentrations in the Eastern United States between 1993 and 2002. They found that the Eastern U.S. could be divided into five distinct regions (e.g. Mid-Atlantic, Great Lakes) that each exhibited distinct temporal patterns (e.g. seasonal trends and persistency) in ozone concentrations. They suggest that their "results suggest that there is a statistically based rationale for delineating geographical areas when interpreting O_3 concentrations" (Lehman *et. al.* 2004, pg. 4368). They propose further work that will categorize the effects of meteorology on ozone concentration in a similar manner.

2.2. Damages from Ozone Pollution and its Precursors

As discussed in Section 2.1, weather conditions and precursor emissions drive ozone formation. These factors can sometimes combine to cause periods of high ozone concentrations. It is these times of high concentration that harm human health. This section reviews the evidence that ozone damages human health and welfare. It also

---

[10] Further research on these issues is out of the scope of this dissertation but will continue under the broader project at the Center for Energy and Environmental Policy Research, which supported this research.

discusses the non-ozone related damages of NO_X and VOC emissions. The argument of some environmental groups that differentiation for NO_X reductions is not worthwhile because NO_X reductions always reduce harmful quantities of some pollutants even if they do not reduce ozone is not always founded. The examples in this section illustrate the tradeoffs that reducing NO_X can create. For example, NO_X reductions that decrease ozone concentrations can simultaneously increase the concentrations of harmful particulate matter in some situations.

2.2.1. Ozone (O_3)

Epidemiological and toxicological studies, including controlled human exposure studies, have linked short-term, or acute, ozone exposure (i.e. exposure that lasts less than 8 hours) to health problems for concentrations of ozone at or above 0.08 ppm (for an extensive literature review see U.S. EPA 2006a). The associated health problems include a reduction in lung function in healthy adults, the onset or aggravation of asthma and of other respiratory diseases, and an increased susceptibility to infections. Some recent controlled human exposure studies have also linked ozone exposure at lower concentrations to respiratory effects in healthy young adults. For example, Adams 2006 found that 6.6 hour-long exposures to 0.06 ppm significantly affected the lung function of health young adults as they exercised (Adams 2006). Epidemiological studies of children and toxicological animal studies show that longer term, or repeated ozone exposure may permanently damage the lungs and cause premature death (see U.S. EPA 2006a, pg. 8-50). Ozone can also reduce agricultural yield, damage tree foliage, and reduce visibility (for a summary see U.S. EPA 2006a, Chapter 9).

Recent results suggest that there is a relationship between exposure to ozone at concentrations observed in many U.S. cities and increased mortality risk. Bell *et al.* 2004 estimated the association between mortality and daily and weekly exposure to ozone in 95 U.S. cities that account for 40% of the U.S. population. Their lagged model estimated the effects of ozone concentrations in a particular day on subsequent days' mortality figures. They also estimated the association between the risk of mortality and to the previous week's cumulative ozone levels. They found, for example, that an increase of 0.01 ppb in ozone from the previous week's concentration was associated with a 0.52% increase in non-injury-related daily mortality and a 0.64% increase in cardiovascular and respiratory mortality (Bell *et al.* 2004). Subsequent studies have supported these findings (e.g. Bell *et al.* 2005, Ito *et al.* 2005, and Levy *et al.* 2005). These results are important because, as discussed later, the association between mortality and pollutant concentrations greatly increases the estimated benefits from reductions in the ambient concentrations of that pollutant in populated areas.

2.2.2. Nitrogen oxides (NO_X)

Beyond their role in the formation of ground level ozone, NO_X emissions cause and contribute to other environmental and human health problems. NO_X emissions contribute to the formation of acid rain and some types of particulate matter (PM) pollution like particulate nitrate (NO_3-). Small, or fine, particulate matter (i.e. particles less than 2.5 μm in diameter) is very detrimental to human health and nitrate is one of these particles, although it is not known whether NO_X-based PM or SO_2-based PM or other species are more harmful than others (see U.S. EPA 2004, Chapter 8). Vehicle catalytic converters turn nitrogen oxides into nitrous oxide (N_2O), which is a greenhouse gas (U.S. EIA 2006, Chapter 4). Nitrogen oxides can directly affect ecosystems through nutrient overloading. For example, the dry and wet deposition of NO_X from the air can cause nitrogen overloading in bodies of water, which accelerates eutrophication and oxygen depletion and harms fish and shellfish populations (see, for example, Ryther and Dunstan 1971). Nitrogen dioxide (NO_2), one of the major components of NO_X also directly affects human respiratory health.

For this last reason and environmental problems like eutrophication, NO_X is itself a criteria pollutant under the Clean Air Act. For the purposes of the NO_X NAAQS, the EPA uses NO_2 as an indicator pollutant for the broader category (in atmospheric chemistry NO_X includes only NO_2 and NO but in other disciplines also includes compounds such as nitric acid (HNO_3) and N_2O). The EPA reviewed the NO_2 NAAQS in 1993 but did not change the standards. In their "National Air Quality and Emissions Trends Report, 2003 Special Studies Edition," the EPA found that all regions of the U.S. were in attainment with the NO_2 NAAQS of 0.053 ppm annual arithmetic mean, in fact most areas of the U.S. had attained the standard by the early 1980s (U.S. EPA 2003a).

The fact that the entire U.S. is in compliance with the NO_X NAAQS means that the secondary pollutants like ozone and PM now drive the implementation of regulations to reduce NO_X emissions. A related issue is that the persistent problems like ozone and PM depend on factors in addition to NO_X emissions and, hence, solving these problems is not as simple as requiring NO_X reductions. For example, ozone formation depends on the chemistry discussed above and reductions in NO_X can sometimes worsen ozone. Similarly, the formation of particulate nitrate depends on NO_X emissions and other factors like the availability of ammonia, moisture, and oxidizing agents. Some studies have found relationships between levels of NO_X and the formation of particulate nitrate that are similar, and related, to those for ozone formation. For example, Pun and Seigneur (2001) found that decreasing NO_X emissions in the winter in the San Joaquin Valley, California could increase the formation of particulate nitrate. Interestingly, in this case, the *increase* in PM came from the *decrease* in ozone caused by the *reduction* in NO_X. Similarly, Blanchard and Tanenbaum (2003a) found that 5 of 16 areas of California had opposite trends in NO_X and particulate nitrate concentrations between 1980 and 2000 (two with increasing nitrate and decreasing NO_X concentrations and three with increasing NO_X and decreasing nitrate concentrations). They also found that lower NO_X emissions on the weekends (the weekend effect) in southern California did not correspond to lower particulate nitrate concentrations (Blanchard and Tanenbaum 2003b).

2.2.3. *Volatile Organic Compounds (VOCs)*

The category of VOC emissions includes a number of different compounds. For example, it includes methane (CH_4), which is a greenhouse gas, other hydrocarbon compounds, and aromatic compounds like benzene and xylene, which are thought to be carcinogens. The hydrocarbon VOCs are those that contribute to the formation of ozone.[11] VOC emissions contribute to water quality issues as well as to indoor air pollution. Products like paints, cleaning supplies, pesticides, and building materials emit VOCs. And concentrations of VOCs are often ten times higher indoors than outside.[12] For the purposes of outdoor air quality, VOCs primarily contribute to the ozone problem.

2.2.4. *Summary*

There are human health and environmental reasons to reduce ozone concentrations. There are also other reasons to reduce NO_X and VOC emissions other than ozone. But the literature discussed in Section 2.1 suggested that if ozone concentrations are to be reduced, regulations must focus on ozone formation chemistry. If this chemistry suggests that NO_X should not be reduced in some areas then tradeoffs between reducing ozone concentrations and other goals should be assessed. However, as discussed later in Section 2.4.3, the statutory framework for air quality in the U.S. makes this a challenge.

---

[11] 40 CFR Part 51.100(s) defines these non-methane VOCs for policy purposes as "any compound of carbon excluding carbon monoxide, carbon dioxide, carbonic acid, metallic carbides or carbonates, and ammonium carbonate, which participates in atmospheric photochemical reactions" except those that appear on a list of exempted compounds. These exempted compounds have "negligible photochemical reactivity."

[12] For a more complete summary see EPA, "An Introduction to Indoor Air Quality: Organic Gases (Volatile Organic Compounds – VOCs)," at http://www.epa.gov/iaq/voc.html, retrieved January 24, 2007.

2.3. Sources of and trends in ozone and precursor emissions

There are three major categories of emissions sources that contribute to atmospheric concentrations of NO_X and VOCs. These are mobile and stationary sources (which are anthropogenic) and natural sources (which are also called biogenic). This section discusses the contributions of each of these source categories to NO_X and VOC emissions in the Eastern United States.[13] It also summarizes available information on the variation of each of these contributions in time and by location and on the recent trends in emissions from the different source categories. Few data are available that describe the variation of mobile source emissions in time and by location. This is one reason why it may be easier to implement a differentiated regulation for stationary sources, for which good data exist. This section also shows that the regulations in the U.S. have reduced ozone precursors significantly and that ozone concentrations have also been reduced. A differentiated regulation could help achieve the final reductions in ozone concentrations required for all areas to meet the air quality. It could also help achieve the reductions that would be required to meet more stringent future standards.

2.3.1. Nitrogen oxides (NO_X)

The major natural sources of NO_X are the burning of biomass (e.g. forest fires), soil release, and lightening discharge (Bond *et al.* 2001 and Zhang *et al.* 2003). These natural sources contribute about 14% of all annual NO_X emissions in the United States (Zhang *et al.* 2003). Although these sources do contribute to the formation of ground-level ozone, they are difficult to control and are dwarfed by the NO_X emissions from the combustion of fossil fuels and other anthropogenic sources.

Mobile anthropogenic sources of NO_X include on-road vehicles (like cars and trucks) and non-road vehicles (like tractors, airport vehicles, airplanes and ships). In

---

[13] The cited EPA documents report these data aggregated for Minnesota, Iowa, Missouri, Arkansas, Louisiana, and the states east of these. These data are available disaggregated by county for years prior to 2002.

2005, the EPA reported that mobile sources contributed 59% of total U.S. anthropogenic NO_X emissions in 2005 (on-road sources like cars and trucks contributed 38%). Stationary anthropogenic sources of NO_X include power plants, industrial boilers, and other industrial facilities. Power plants and other large industrial sources contributed 22% of the total 2005 NO_X emissions in Eastern States (U.S. EPA 2006b), with about 97% of this contribution from power plants.[14]

The time and locational aspects of ozone chemistry suggest that studying annual trends in NO_X, VOCs, and ozone for the large area of the entire Eastern U.S. may not be the most useful way to understand the problem. Unfortunately, very few detailed datasets are available that would enable location-specific analyses of trends in NO_X emissions for daily or hourly periods. While the Continuous Emissions Monitoring System (CEMS) provides detailed, hourly data on NO_X emissions from power plants and other large industrial sources, there are not similar data available for mobile sources.[15] Likewise, the CEMS data provide information on individual generating units and industrial sites, but data for mobile sources from particular locations are not widely available. The difficulties involved in directly monitoring emissions from mobile sources are likely the reason for this dearth. EPA does publish estimated annual emissions data at the state and county level for both mobile and stationary sources and for designated "statistical metropolitan areas", but the most recent year of data that are easily queried is 2001 and the EPA has only published data up through 2002.[16]

Figure 2-1 shows the trends in the shares of stationary and mobile source emissions of the annual NO_X emissions for selected Eastern States between 1999 and

---

[14] This latter fact was calculated from EPA Continuous Emissions Monitoring data at http://cfpub.epa.gov/gdm/.

[15] The CEMS data are available at the EPA's "Clean Air Markets – Data and Maps" at http://cfpub.epa.gov/gdm/ under "Emissions."

[16] Data through 2001 are available at the EPA's "AirData" at http://www.epa.gov/air/data/index.html. Data through 2002 are available at the EPA's "2002 National Emissions Inventory Data & Documentation" at http://www.epa.gov/ttn/chief/net/2002inventory.html.

2001.[17] The figure suggests that the relative contributions of different sources to total NO_X vary geographically; stationary sources contributed between 5 and 45% of states' total annual NO_X emissions in this period. At the county level, the variation was greater. On average, stationary sources contributed about 15% of county-level NO_X emissions in Eastern States in 2001. They contributed over 50% of NO_X emissions in about 10% of eastern counties and over 90% of NO_X emissions in about 1% of counties.[18]

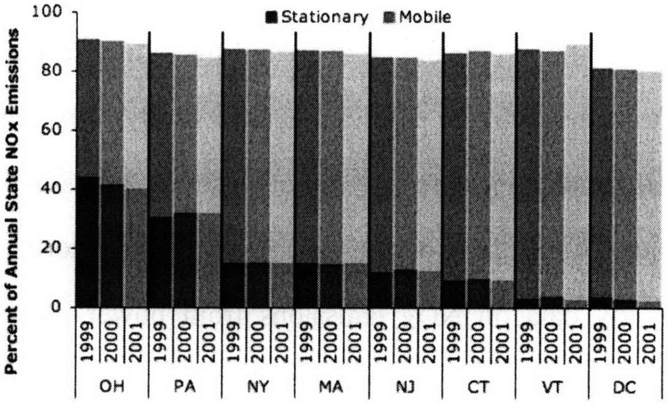

Figure 2-1 Stationary and mobile source shares (percent) of total, annual NO_X emissions in some Eastern States for the years 1999, 2000 and 2001.

Although the Figure 2-1 data are not yet publicly available for 2005, the EPA has summarized these data in a report on the effectiveness of NO_X regulations. They reported that stationary sources contributed about 22% and mobile sources about 59% of NO_X emissions in all Eastern States in 2005 (U.S. EPA 2006b). In the same Eastern States in 2001, stationary sources contributed about 33% and mobile sources about 54% of total NO_X emissions.[19] This change in the relative contributions of stationary and mobile

---

[17] Data are from the EPA's "AirData" at http://www.epa.gov/air/data/index.html.

[18] County-level calculations were also performed with data from the EPA's "AirData" at http://www.epa.gov/air/data/index.html.

[19] The 2001 percentages were calculated from the EPA's "AirData" at http://www.epa.gov/air/data/index.html using the Tier-1 emission categories of fuel combustion for electricity generation and other industrial processes for stationary source emissions and the Tier-1 categories of on- and off-highway vehicles for mobile sources. In all cases, "Eastern states" refers to Minnesota, Iowa, Missouri, Arkansas, Louisiana, and the states east of these.

sources was partially a result of major control programs that the EPA and state regulators implemented for NO_X from stationary sources between 1999 and 2006 (see Chapter 4).

The stationary-source shares of NO_X emissions also decreased between 1999 and 2005 because NO_X emissions from mobile sources increased. The EPA's regulations requiring mobile sources to reduce NO_X emissions with both fuel and technology standards have reduced emissions per mile from vehicles, but increases in the use of non-road diesel engines and in vehicles miles traveled for both heavy-duty diesel and gasoline trucks more than offset the emissions reductions from these programs in many parts of the United States (U.S. EPA 2003a). The large decreases in emissions from stationary sources like power plants have offset this increase and caused total annual NO_X emissions to decrease; the EPA estimates that total NO_X emissions in the Eastern U.S. have fallen from about 16 million tons annually in 1999 to about 12.5 million tons in 2005 (U.S. EPA 2003a and 2006b).

In addition to the geographic dependence of relative contributions of mobile and stationary sources to total NO_X emissions, human activity patterns probably cause diurnal and seasonal variations as well. For example, the contribution of power plants to NO_X total emissions in some populated areas is likely higher at night, when there is less traffic than during the day. Unfortunately, it is difficult to find data on mobile source emissions at the frequency needed to analyze these trends – although such information could ultimately contribute to a more cost-effective approach to regulation if a careful monitoring of these trends could help identify the areas and times that particular sources contributed most to local and even regional, ozone formation.

2.3.2. Volatile Organic Compounds (VOCs)

In contrast to NO_X emissions, natural sources of VOC emissions contribute substantially to total VOC emissions in many parts of the United States. Examples of biogenic VOC (BVOC) emissions are isoprene and terpenoid emissions from trees and other plants. Chameides *et al.* (1988) found that 60% of Atlanta's 11-county urban area was wooded and that the daily emission rate of BVOCs in this area was greater than that from anthropogenic sources (pg. 241, note 13). Pierce *et al.* (1998) confirmed findings since 1988 that BVOC emissions equal or slightly exceed those from anthropogenic sources in the Eastern United States; they back-calculated BVOC emissions for periods when scientists had measured BVOCs. Later studies have also confirmed these findings and studied how changing land-use patterns have caused BVOC emissions to shift across the Eastern U.S. over time (Purves *et al.* 2004, Fiore *et al.* 2005).

Key to these estimates of BVOC emissions was the finding that the quantities of some BVOCs emitted from plants are a function of sunlight or heat. Lamb *et al.* (1987) found that a rise in ambient temperature from 25 to 35 degrees Celsius caused some deciduous trees to increase their isoprene emission rate by a factor of four and some conifers to increase their terpene emission rate by a factor of one and a half. In general, studies have found that BVOC emission rates increase exponentially with the temperature of trees' leaves up to about 40 degrees Celsius (Tingey *et al.* 1979 and 1980). Isoprene emissions from some plant species are a by-product of photosynthesis and therefore depend on the presence of sunlight; isoprene emission rates increase with increasing light intensity (NRC 1991, pg. 263). The observation that trees and other plants emit more BVOC emissions on hot, sunny days is important for policy as these days are also the most conducive to the formation of ozone.

Biogenic VOC emissions, especially highly reactive isoprene, play an important role in the formation of ground-level ozone in the Eastern United States and are relevant to air quality policy (Trainer *et al.* 1987). Pierce *et al.* (1998) found that the addition of biogenic VOCs to air chemistry models caused a substantial increase in predicted ozone levels and that accounting for BVOC emissions categorized many areas of the Eastern United States as NO_X-limited, instead of VOC-limited. This meant that reductions in

NO_X emissions were, and still are, more crucial for mitigating ozone formation than reductions in anthropogenic VOC emissions in the Eastern United States (see, for example, McKeen *et al.* 1991 and Sillman *et al.* 1990). High levels of isoprene emissions can also reduce ozone when NO_X emissions are low (Kang *et al.* 2003).

Sources of anthropogenic VOC (AVOC) emissions include combustion processes, dry cleaning, and fumes from substances like solvents, fossil fuels, and paints. Solvents contributed 27% and mobile sources 39% of AVOC emissions in the Eastern U.S. in 2005 (U.S. EPA 2006b). Solvents and paints emit AVOCs through evaporative processes, as do gasoline and other petroleum products. Mobile sources like light-duty and heavy-duty trucks and lawnmowers emit AVOCs when fuels permeate through hoses and fittings and during start-up and operation.

Between 1990 and 2005, AVOC emissions decreased from about 16 million tons to about 11 million tons per year in the Eastern United States (U.S. EPA 2006b). Decreases in VOC emissions from the transportation sector caused the majority of this reduction (U.S. EPA 2003a). There is little information on the diurnal and seasonal patterns of AVOC emissions; in fact, uncertainties around VOC emissions are greater than for NO_X because large stationary sources are responsible for a large portion of NO_X emissions. Emissions from sources like power plants are easier to estimate or monitor, estimation procedures for evaporative emissions and mobile source emissions, the latter also affect NO_X estimates, are more difficult (see for example, NRC 1991, pg. 252-4).

2.3.3. Ozone (O_3)

There are few direct anthropomorphic sources of ozone pollution, the only major one being ozone from outside the North American boundary layer. The lifetime of ozone is about one week, which is long enough for it to be transported between continents

(Fiore *et. al* 2002).[20] Scientists and policymakers in the U.S. consider this contribution part of the policy-relevant ozone background concentration. Fiore *et al.* (2002) found that background ozone from outside the North American boundary layer contributed about 0.015 ppm in the Eastern United States during the stagnant, summertime conditions most conducive to the formation of high ozone concentrations, and that it contributed more at other times (Fiore *et al.* 2002).[21]

Ozone concentrations have decreased since the 1980s in the Eastern United States, but the decreasing trend slowed after 1990. In the late 1980s and early 1990s, studies in the literature indicated that BVOC emissions caused NO_X-limited conditions for ozone formation in the Eastern United States, and regulators realized that they needed to control NO_X as well as VOC emissions (U.S. EPA 2006b). The EPA and the eastern states implemented major regulations between 1996 and 2006 that controlled NO_X emissions from large stationary sources, mostly power plants (Chapter 4).

To measure general trends in ozone concentrations, the EPA calculates the "seasonal average 8-hour ozone concentrations" (U.S. EPA 2006b). They average the daily maximum 8-hour ozone concentrations from each monitor from May 1st through September 30th each year, while accounting for differences in meteorology between years. In eastern states, the average reduction in ozone concentrations was between 6 and 21% from 1983 to 2002; the largest decrease occurred in the Northeast and the smallest in the Midwest. With the exception of the northeastern states, little or none of this reduction occurred between 1993 and 2002 (U.S. EPA 2003a). Between 2002 and 2004, the average reduction was 8% in all eastern states and there was no reduction in average ozone concentrations between 2004 and 2005 (U.S. EPA 2006b). In 2005 the EPA found

---

[20] The lifetime of ozone over the Eastern United States in particular is shorter, usually about 2 days, because of shallow mixing depth, interactions with biogenic VOCs, and other factors (Fiore *et al.* 2002).

[21] The EPA found the policy-relevant background levels of ozone in the U.S. in general to be about 35 ppb, but noted that they do vary in time and space (EPA 2006a, Chapter 3, pages 44-55).

that the seasonal average 8-hour ozone concentration in the Eastern United States was about 0.053 ppm (U.S. EPA 2006b). These changes, especially the recent ones, have improved local air quality in many areas. The following sections of this chapter discuss these changes in local air quality in more detail, as well as their relationship to the federal air quality standards for ozone.

2.4. Ozone National Ambient Air Quality Standards (NAAQS)

The first part of this section briefly reviews the history of the federal air quality standards for ozone. Section 2.4.2 then uses ozone monitoring data to illustrate the difficulties that states in the Eastern U.S. have had attaining these standards. The persistent non-attainment of the standards, despite the reductions in precursor emissions just discussed in Section 2.3, provides motivation for the consideration of a differentiated approach to the resolution of NO_X to control ozone.

Section 2.4.3 builds of the history of the history of U.S. policy and discusses why the statutory frameworks for air quality require a focus on achieving the air quality standards rather than on balancing the costs and benefits of NO_X emission reductions. The air quality management approach mandated by the Clean Air Act requires the EPA to set air quality standards based on science without reference to the cost of achieving them. The EPA and state governments use regulations to achieve the standards through reductions in precursor emissions. The formation chemistry of secondary pollutants like ozone creates a challenge for this approach because straightforward reductions in precursor emissions do not always cause decreases in the concentrations of the secondary pollutants and reductions in precursor emissions can lead to tradeoffs. Differentiated regulations working within the statutory framework may help mitigate this problem.

2.4.1. History of the National Ambient Air Quality Standards

In large part due to its complicated chemistry, ground-level ozone pollution has been a persistent environmental problem and a policy, technical, and scientific challenge since urban air pollution caught the public's attention in the 1940s. Congress officially recognized ground-level ozone and other photochemical oxidants as a widespread problem in 1970 when it categorized them as one of six "criteria pollutants".[22] The Clean Air Act of 1970 (CAA) mandates the Environmental Protection Agency (EPA) to set two science-based National Ambient Air Quality Standards (NAAQS) for criteria pollutants.[23] It requires "primary" standards to protect public health with an "adequate margin of safety" and "secondary standards" to protect against other welfare effects like those on ecosystems and visibility.[24] The CAA mandates the EPA to periodically review the NAAQS.[25] It also requires states to develop State Implementation Plans (SIPs) to regulate various emission sources to levels that would ensure attainment of the NAAQS.[26]

The EPA first set a standard for photochemical oxidants in 1971.[27] Ozone is a photochemical oxidant and typically makes up over 90% of all photochemical oxidant pollution (see, for example, Bates 1983). The EPA reviewed the photochemical oxidant standards between 1976 and January 1979. Based on a review of scientific studies of exposure, health, and other environmental impacts of these pollutants, the EPA revised the photochemical oxidant NAAQS by designating ozone (O_3) as the indicator for the category of pollutants. This means that the EPA felt it was more appropriate to measure and regulate only ozone concentrations, rather than doing so for all the pollutants in the

---

[22] Congress, in CAA Sections 108 and 109, required the EPA to identify and list criteria pollutants: those that are widespread problems and emitted by many industrial sectors, and that have a significant effect on public health and welfare. The six criteria pollutants are NO_X, ozone, sulfur dioxide, lead, carbon monoxide, and particulate matter.

[23] CAA §108(a)(2) states: "Air quality criteria for an air pollutant shall accurately reflect the latest scientific knowledge useful in indicating the kind and extent of all identifiable effects on public health or welfare which may be expected from the presence of such pollutant in the ambient air, in varying quantities."

[24] This is true for all criteria pollutants under the Clean Air Act, Section 109(a), 42 U.S.C. § 7409.

[25] Section 109(d) of the CAA (42 U.S.C. 7409) mandates five-year reviews.

[26] CAA Section 111(d).

[27] The standard was that the daily maximum one-hour average could not exceed 0.08 ppm more than one day per year.

category, as they found ozone to be the most important and representative of the broader category. The 1979 ozone NAAQS required the daily maximum hourly average ozone concentrations not to exceed 0.12 parts per million for more than three days over a four year period.[28] This standard is often called the "1-hour" standard. The EPA updated the ozone standards in 1997.[29] The 1997 ozone NAAQS is the current standard and it requires the 3-year average of the fourth-highest daily maximum 8-hour average ozone concentrations each year not to exceed 0.08 parts per million. This is typically referred to as the "8-hour" standard. The EPA is currently reviewing the 8-hour standard and the EPA staff scientists have recommended that the Administrator increase its stringency to within the range of 0.06 to 0.08 ppm (U.S. EPA 2007b).

2.4.2. Areas and times of persistent non-attainment of the ozone NAAQS

Many counties of the Eastern U.S. are struggling to meet the 8-hour ozone NAAQS and an increase in stringency of the standard would intensify this challenge. In 2004, the EPA used 2001 through 2003 data to designate 103 areas as nonattainment areas for the 8-hour standard.[30] In a 2006 report, however, the EPA found that, based on 2003 through 2005 data, only 31 of these areas were still in nonattainment (Figure 2-2, U.S. EPA 2006b). About 81 million people live in these 31 remaining nonattainment areas, which is roughly 40% of the total population in the Eastern United States.[31] These areas' high populations mean that it is important from a public health perspective for them to attain the NAAQS for ozone.

---

[28] The three days in four years part of this standard ensures that the annual "expected number" of exceedences will less than one. If three or fewer exceedences occur in four years then the average number of exceedences over those four years will be less than one. In this way the EPA approximated a long-term average, rather than penalizing states with one or two "outlier" exceedences over four years.

[29] Federal Register, Vol. 62, No. 138, Friday, July 18, 1997.

[30] The EPA designated 126 areas in the entire U.S. as nonattainment areas at this time. See U.S. EPA, "Air Quality Designations and Classification for the 8-Hour Ozone National Ambient Air Quality Standards (NAAQS)" *Federal Register* 71(93), May 15, 2006.

[31] The EPA defines the "Eastern" states as Minnesota, Iowa, Missouri, Arkansas, Louisiana, and the states east of those (see, for example, U.S. EPA 2006b). The U.S. Census Bureau estimated that the population in these states in July 2006 was about 197 million (calculated from U.S. Census Bureau, "National and State Population Estimates" at http://www.census.gov/popest/states/NST-ann-est.html).

Figure 2-2 Nonattainment areas for the 8-hour ozone standard. Calculated by the EPA with 2003 through 2005 data (Figure from U.S. EPA 2006b).

Data from the EPA's air quality monitoring database suggest that the ozone concentrations in nonattainment areas only exceeded the 8-hour standards on a few days each year. For example, in the Northeastern and Mid-Atlantic states, counties in nonattainment areas violated the 8-hour standard on about 4 days per year on average between 2003 and 2005.[32] The county with the most daily exceedances during this three-year period was Ocean County, New Jersey. Ozone readings from this county's single monitor exceeded the 8-hour standard on 30 days during the three-year period (9, 7, and 14 times in 2003 through 2005 respectively). The monitors in the broader Philadelphia-Wilmin-Atlantic "moderate" nonattainment area, which includes Ocean County, recorded an average of 3.6 days per year that exceeded the 8-hour standard.[33] The area's

---

[32] Figures calculated from data retrieved from the EPA's AirData website (http://www.epa.gov/air/data/index.html). Data for ozone air quality monitors in the states of EPA Regions 1, 2, and 3 (CT, DC, DE, MA, MD, ME, NH, NJ, NY, PA, PR, RI, VA, VI, VT, WV) was aggregated to the county-level by taking the average of the 4th highest 8-hour ozone reading for all monitors in each county for the three years. Nonattainment counties were assumed to be those with values greater than 0.08 ppm for this calculation. (It is worth noting that the averaging across monitors had little effect, because the majority of counties only reported data from one monitor.) The query from the EPA's database also returned the number of days in each year that each monitor exceeded the 8-hour standard. These data were summed across all years for each monitor and then the total 3-year exceedances were averaged across monitors in each county.

[33] For more details on this nonattainment area and the counties it includes (and on other nonattainment areas), see the EPA's "Greenbook" at http://www.epa.gov/oar/oaqps/greenbk/gnca.html. The data retrieved from the EPA's "AirData" included those for 29 ozone air quality monitors in the counties in this nonattainment area; these reported 85 monitor-years of data between 2003 and 2005 and 307 monitor-days that exceeded the standard.

population is over 7 million people and "moderate" is the worst level of nonattainment in the Eastern U.S. currently, areas of California suffer "serious" or "severe" nonattainment.

The few number of days in which the worst nonattainment areas in the Eastern U.S. suffer from ozone concentrations above the standard suggests that if additional undifferentiated regulations are implemented to address this problem, the potential for costly over-compliance is high. On many days of the summer, additional abatement is not necessary. The role of this research is to determine the feasibility of linking forecasts of the critical days to market-based incentives for NO_X abatement from sources that impact ozone concentrations in these targeted areas.

2.4.3. Choosing acceptable ozone concentrations

In a report on the ozone problem, the National Research Council referred to the NAAQS approach – setting ambient air quality standards and then designing regulations that reduce emissions until the ambient standards are met – as the "air quality management approach" to environmental regulation (NRC 1991, pg. 251). They noted that this strategy can be effectively implemented for primary pollutants; for example, carbon monoxide (CO) emissions lead directly to elevated ambient concentrations of CO that harm human health and, thus, reducing CO emissions solves the problem. But they also discuss why what they call the "air quality management approach" is difficult to apply successfully to secondary pollutants: There is not a direct linear, casual relationship precursor emissions and the formation of most secondary pollutant, like ozone and particulate matter, so reducing those precursor emissions is not a guaranteed strategy to mitigate the air quality problem.

Both ozone and particulate matter (PM) pollution provide an example of this challenge. Reductions in NO_X emissions do not always lead to decreases in ozone and PM. In some cases, like those discussed in Section 2.2.2, the same reduction of NO_X can lead to decreases in one pollutant and increases in the other. This tradeoff illustrates a major challenge for what the NRC calls the "air quality management approach" under the statutory framework of the Clean Air Act (CAA). Ideally tradeoffs like this would be resolved by comparing the costs of incremental reductions in emissions to the total marginal benefit of further reductions in emissions. Then if NO_X reductions were worthwhile, given the tradeoffs, they could be undertaken.

This presents problem because the EPA is not legally able to consider the costs of setting the NAAQS; their goal must only be to protect the public health with an "adequate margin of safety." [34] The literature discusses the problems associated with the EPA's inability to consider costs while setting the NAAQS: its inability to consider cost prevents the EPA from using one of the tools available to determine the level of acceptable risk. Ozone, for example, is not a "threshold" pollutant.[35] This means that science, as of yet, has not established levels below which exposure to ozone is safe so the EPA must choose a standard that is based on the available science. Once this air quality standard is established it must be met, even if the emission reductions required to meet it cause increases in other harmful pollutants.[36] Coglianese and Marchant (2004) argued that the EPA's inability to consider costs has allowed them to avoid providing internally consistent and transparently reasoned justifications for the NAAQS decisions (pg. 1292). They suggest the EPA could more effectively protect the public health by focusing resources on reducing particulate matter pollution rather than ozone.[37] They argue more

---

[34] This is true for all criteria pollutants under the Clean Air Act, Section 109(b)(1), 42 U.S.C. § 7409(b)(1). In *Whitman v. American Trucking Associations*, the Supreme Court found that "The text of §109(b), interpreted in its statutory and historical context and with appreciation for its importance to the CAA as a whole, unambiguously bars cost considerations from the NAAQS-setting process..." 531 U.S. 457 (2001).

[35] Limitations of epidemiological research and the lack of available data at low exposure levels make it difficult to detect a this type of threshold and, at this stage, there is not conclusive evidence as to whether or not one exists for ozone (EPA 2006a, Chapter 7, pgs. 154-9). The literature often models the damage function of ozone as linear or log-linear based on a linear or log-linear concentration-response function without a threshold below which exposure is safe (see, for example, Tong *et. al.* 2006 using the concentration-response function estimated in Bell *et. al.* 2004).

[36] Another problem is that as science continues to improve and scientists are able to identify the effects of pollutants on human health at lower and lower concentrations, the EPA will not be able to eliminate all risks associated with exposure to very low concentrations without imposing extremely high costs on society.[36] The EPA has recognized that zero-risk standards are not necessary or desirable.[36]

[37] "In refusing a more stringent alternative for the PM standard, EPA rejected an option that would have achieved a much greater gain in health benefits than the gain EPA anticipated from its revision of the ozone standard. If protecting

generally that by comparing the costs and benefits of achieving various levels of ambient standards for each criteria pollutant, the EPA could evaluate whether it was more cost effective to concentrate on reducing ozone or other pollutants like particulate matter (Coglianese and Marchant 2004, pgs. 1334-5).

The use of differentiated regulations to address secondary pollutants with important spatial and temporal relationships could mitigate the challenge of accounting for tradeoffs between environmental problems while not considering costs and benefits. A differentiated approach for both ozone and PM, for example, would at least enable regulators to explicitly address the tradeoff involved with reductions in NO_X even if they could not balance the costs and benefits of reducing ozone compared to PM (many areas unable to attain the ozone standard also cannot attain the PM standard). Differentiation could enable regulators to move away from their current strategy of simply requiring reductions in precursor emissions.

A possible extension of the research in this dissertation, and of the broader research program into which it fits, is to estimate the marginal costs and benefits of reducing ozone concentrations (and even PM concentrations) in the Eastern U.S. through NO_X reductions and thus to eventually inform policy decisions about efficient levels of ambient concentrations of ozone and other pollutants. For now, however, this research takes as given the goal of achieving the ozone NAAQS. Studying how remaining ozone nonattainment areas might achieve compliance cost-effectively is not at odds with the goal of understanding the marginal costs and benefits of reducing NO_X and ozone. The remaining nonattainment areas are highly populated and, thus, hold the greatest potential for human health benefits from improved air quality. Also, this dissertation seeks to understand one of the potential inputs to a full benefit-cost analysis: could power plant

---

the public health with an adequate margin of safety did not require the Agency to lower the PM standard still further, then it is far from clear why the Agency was justified in revising its ozone standard at all," (Coglianese and Marchant 2004, pgs. 1321-22).

dispatch provide flexible NO_X abatement options that could be targeted to reduce the probability of high ozone concentrations in populated areas?[38]

2.5. The regulation of ozone precursor emissions

While the CAA and its required State Implementation Plans (SIPs) mean that states are largely responsible for developing and implementing their own regulations to meet the ozone NAAQS, the EPA has driven most of the major initiatives to reduce the emissions of ozone precursors. The EPA has aimed some of these actions at helping states achieve the ozone air quality standard, and others at different problems like acid rain and particulate matter. This section discusses federal mobile and stationary source regulations for NO_X and VOCs. The history of federal regulations shows that the regulations have evolved to some extent with the scientific understanding of the relationships between precursor emissions and ozone concentrations. But even with the planned reductions in NO_X emissions under the most recent regulations, the EPA does not expect all areas of the Eastern U.S. to attain the ozone air quality standards. This suggests that more attention to the science of ozone formation is needed.

2.5.1. Mobile source regulations

The history of mobile source regulations in the U.S. reaches back to the 1950s when California researchers first recognized the connection between vehicle emissions and the formation of photochemical smog, which consists of both ozone and particulate matter (NRC 2004). California led the country in its adoption of emission standards for new vehicles starting with the 1966 model year; the federal government followed suit with the Motor Vehicle Pollution Control Act that implemented national vehicle emission

---

[38] Further related research questions are: Could other sources such as vehicles generate targeted NO_X reductions at a lower cost? (Chapter 6 briefly discusses this.) Would targeted actions be more cost-effective than blunt actions like further decreases in summertime emissions caps? Given the costs of these options to reduce NO_X and the reductions in ozone they are expected to cause, what would be the benefits of implementing them? (Section 3.2.4 reviews studies that have estimated the costs and benefits of NO_X reductions.)

standards starting in the 1968 model year.[39] Since this time, regulators have used four mechanisms to limit emissions from mobile sources: emission standards (in grams per mile) for new vehicles and motors, fuel property specifications, in-use vehicle inspection and maintenance programs, and incentives for behavioral changes (e.g. transportation management programs). Regulations based on these mechanisms have achieved mixed success.

The success of each of these four mechanisms is limited, at least to some degree, by whether the others are implemented. For example, political challenges have constrained the use of incentives to reduce driving, especially in metropolitan nonattainment areas.[40] Vehicle miles traveled (VMT) increased by about 150% between 1970 and 2003 (U.S. EPA 2003a). Thus, although new vehicle emission standards have encouraged the development of technologies to drastically reduce vehicle tailpipe emissions (Table 2-1), the increases in VMT and inability of regulators to influence driving behavior have offset at least some of the progress made in reducing emissions-per-mile. In fact, prior to 1994 aggregate NO_X emissions from light-duty gasoline cars and trucks increased (U.S. EPA 2003a, pg. 19). In addition, a disproportionately large fraction of emissions come from a small number of older vehicles and vehicles with dysfunctional emissions control equipment.[41] Inspection and maintenance programs have not been able to eliminate this problem, again limiting the ability of new vehicle standards to reduce aggregate emissions (NRC 2001).

---

[39] California's mobile source standards have typically led those adopted by the federal government by about two years. The CAA does not allow states other than California to independently set mobile sources standards. Other states can, however, adopt the California standards. For further discussion of this issue see, for example, NRC 2004, pg. 136.

[40] The 1970 Clean Air Act gave the EPA the authority to encourage states to develop transportation control plants (TCPs) in nonattainment areas as part of their SIPs. TCPs might include policies like taxes o surcharges on parking downtown in metropolitan areas or access restrictions or congestion charges in downtown or polluted areas. The CAAA of 1977 and 1990 also gave states the option to use these mechanisms but the states have not chooses to adopt them. An example the political infeasibility of using these controls is the EPA's promulgation of TCPs for 19 metropolitan nonattainment areas in 1973. The states resisted these policies and Congress subsequently restricted the EPA's ability to use price incentives or to restrict parking in this way. See NRC 2004, Chapter 4 for further discussion of these issues.

[41] NRC 2001 summarizes studies that found, for example, that 5% of California passenger vehicles contributed about 85% of NO_X emissions in 1999 (pg. 35).

Table 2-1 U.S. federal emission standards for NO_X and VOC emissions from light-duty and heavy-duty vehicles between 1968 and 2009. Sources: NRC 2004, 2006, U.S. EPA 2000a and 2000b.

Federal Standards	**Passenger Vehicles (Light-duty Vehicles)**		**Heavy-duty Diesel Engines (>8500 lb)**	
Model Year	**VOCs**	**NOx**	**VOCs**	**NOx**
Uncontrolled	8.7	3.4		
1968	4.1	3.4		
...				
1972	3.0	3.4		
...				
1975	1.5	3.1		
...				
1977	1.5	2.0		
...				
1980	0.41	2.0		
1981		1.0		
...				
1993				5.0
1994			1.3	
1995				
1996				
1997				4.0
1998				
1999				
2000				
2001	0.075*	0.2-0.4**		2.0
2002				
2003				
2004		0.3	2.4^	
2005				
2006				
2007		0.07*	0.14	0.2
2008				
2009	0.015	0.07*^		
	g/m		*g/bhp-hr*	

* Fleet average, passenger cars and light-duty trucks
** Emissions standard varies depending on certification level
^ Combined nonmethane VOC + NOx standard
*^ Fleet average, includes medium-duty vehicles (e.g. SUVs)

Regulators have been better able to manage the interdependence between fuel composition and the technologies required to meet new vehicle tailpipe standards. Some substances in fuel – especially tetraethyl lead and sulfur – damage catalytic converters, which is a primary technology used to reduce NO_X tailpipe emissions from both light-duty gasoline vehicles and trucks (NRC 2004, pg. 155). Regulations in 1973 required the phase-out of leaded gasoline thereby mitigating exposure to airborne lead oxides and enabling the use of catalytic converters to reduce NO_X and PM.[42] The EPA banned leaded

---

[42] 38 Federal Register 1255, Jan. 10, 1973, as amended at 38 Federal Register 33741, Dec. 6, 1973

fuels in 1996.[43] The "Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements" reduced the level of sulfur allowed in gasoline starting in 2004 in order to improve the effectiveness of catalytic converters and to prevent the degradation of these technologies over time.[44] The latter is especially important given the limited effectiveness of inspection and maintenance programs.

However, gasoline composition regulations have not been altogether free of problems. For example, refiners replaced the octane enhancing qualities of tetraethyl lead in gasoline by blending higher amounts of light hydrocarbons and aromatics like benzene into the fuel.[45] This increased evaporative VOC emissions and air toxic emissions from gasoline. In the early 1970s California imposed limits on the volatility of gasoline (via the measure of Reid vapor pressure or "RVP") to reduce evaporative VOC emissions from vehicles, storage tanks and distribution facilities. In 1989 the federal government also imposed RVP limits. Then, in the 1990 CAAA Congress required the use of reformulated gasoline during the summer in nine metropolitan, ozone nonattainment areas.[46] Reformulated gasoline generally contains fewer aromatics (like benzene), less sulfur, and has a lower RVP. It also has additional oxygen content compared to regular gasoline, which reduces CO emissions from combustion.[47] The two additives most used to meet the oxygen-content requirements are MTBE[48] and ethanol; both of these additives cause ancillary problems. MTBE contaminates groundwater. Both may cause increases in NO_X emissions (NRC 1999). Ethanol also increases the RVP of gasoline, thereby

---

[43] 61 Federal Register 3832 (February 2, 1996).

[44] See page 6702 in 65 Federal Register 6698 (February 10, 2000).

[45] See NRC 2004, pg. 155. Octane helps gasoline resist knocking while burning in the combustion chamber, for more details see Chevron 2007, Chapter 2.

[46] As of 2004, the EPA required RFG in 10 metropolitan areas in the summer: Los Angeles, San Diego, Baltimore-Washington, Hartford, New York, Philadelphia, Chicago, Milwaukee, and Houston, and Sacramento. They added Sacramento in 1995 when it was reclassified as a severe nonattainment area. Fourteen metropolitan areas or states participate voluntarily or have their own RFG programs (NRC 2004, pg. 160).

[47] See NRC 2004, pgs. 155-157. The incomplete combustion of carbon in fuel results in carbon monoxide emissions and higher oxygen content helps mitigate this problem, see Chevron 2007, Chapter 2.

[48] MTBE is Methyl Tertiary Butyl Ether.

increasing evaporative VOC emissions or canceling the other measures taken to reduce the RVP of reformulated gasoline.[49]

The popularity of medium-duty passenger vehicles (e.g. SUVs) that have been exempt from the emissions standards placed on light-duty passenger vehicles and trucks, increases in the use of nonroad diesel engines, and increases in the VMT of heavy-duty diesel trucks have all contributed to the increases in NO_X emissions from mobile sources since the 1970s (U.S. EPA 2003a). These trends have caused regulators to shift their focus from light-duty passenger vehicles to medium-duty passenger vehicles and nonroad and heavy-duty diesel trucks and motors. For example, the Tier 2 standards for passenger vehicles, which will be fully implemented by 2009, now, for the first time, include medium-duty vehicles like SUVs and vans (U.S. EPA 2000a). The EPA has also established categories of nonroad engines and emission standards for each category that have been phased-in starting in 2000.[50]

The EPA has also increased the stringency of emission standards for heavy-duty diesel trucks in recent years (Table 2-1); these new standards will require significant changes in compliance strategies for diesel truck and engine manufacturers and will be phased-in between 2007 and 2010. These NO_X standards will likely require heavy-duty diesel trucks to adopt technologies like selective catalytic reduction (SCR) with ammonia or NO_X absorber catalysts (NRC 2006, pg. 129). Until now, trucks have been able to meet the NO_X emission standards through engine modifications.[51] In addition to the

---

[49] Many have discussed the controversy over whether political goals to create a subsidy for corn farmers was the major motivation for the oxygenate requirements, rather than air quality. For a further discussion of this issue and of the impact of the use of MTBE and ethanol on VOC and NO_X emissions, see U.S. EPA 1999 and NRC 1999, 2004.

[50] See 63 *Federal Register* 56968 (October 23, 1998) and for a summary see EPA 2003b and DieselNet, "Nonraod Diesel Engines," at http://www.dieselnet.com/standards/us/offroad.html. Also see 67 *Federal Register* 68241 (November 8, 2002) and 68 *Federal Register* 28327 (May 23, 2003).

[51] In fact, because of difficulties in monitoring or testing emissions from trucks under typical driving conditions, the standards have applied to engines only, which were tested in laboratories (thus the units of grams per horsepower-hour rather than grams per mile, as for vehicles). This allowed engine manufacturers to subvert the standards by using software programs in trucks that reverted to operating with better fuel economy and higher NO_X emissions at cruising speeds. This issue was resolved in 1998 (see, for example, NRC 2006 pg. 234).

heavy-duty diesel standards, the EPA has also required a reduction in the sulfur-content of diesel fuels to improve the functioning of NO_X control equipment for diesel engines; sulfur can damage the control technologies needed that reduce NO_X emissions.[52]

One aspect of the history of mobile source regulations that is relevant to this research is the difficulties that regulators have faced when trying to use incentives to alter personal behavior. Short-term, targeted reductions of NO_X emissions in nonattainment areas could be achieved by limiting driving during hours when NO_X reductions are most critical for mitigating ozone formation in highly populated areas. Regulators implemented a program in London that charges fees for driving into central parts of the city. Notably, however, the fees in London are close to those already paid, for example, to cross the Hudson River into New York City; this suggests that much higher fees might be necessary to deter driving in U.S. cities (NRC 2004, pg. 163).

These types of programs may prove more feasible for diesel trucks than for passenger vehicles. Regulators could use price disincentives to limit driving on critical days or during critical times for diesel trucks, or they could mandate the use of controls for the most critical days or hours and relax that mandate during other times. The selective catalytic converters and absorbers used to limit NO_X emissions from heavy-duty diesel trucks have high variable costs. So cost savings from not requiring the constant use of these technologies could be significant and could also extend the life of the control equipment.

2.5.2. Federal stationary source regulations

As the previous section discussed, the EPA's approach to reducing ozone concentrations first focused on VOC emissions, mostly from mobile sources. Then as the deceasing trends in ozone concentrations slowed and science stressed the importance of

---

[52] EPA published the low-sulfur diesel rule in 71 Federal Register 25706 (May 1, 2006). They published the diesel emission standards in 65 Federal Register 59896 (October 6, 2000).

NO_X reductions, especially in the Eastern U.S., the EPA's focus shifted to controlling NO_X emissions from stationary sources (U.S. EPA 2005). Prior to the 1990 CAAA, the federal government placed very few constraints on NO_X and VOC emissions from stationary sources. The 1990 CAAA initiated a cascade of regulations.

Prior to 1990, emission limits for stationary sources applied only to new sources and those undergoing substantial modification. Title I of the 1970 CAAA required the EPA to promulgate New Source Performance Standards (NSPS) for stationary sources. NSPS are emission standards (in units of emissions per heat input, e.g. lbs/mmBTU) that are based on what sources' emission rates would be if they installed the Best Available Control Technology (BACT).[53] The provisions required new sources to undergo a permitting process showing that they would achieve these emission rates and required existing sources to undergo this process when they undertook significant modifications.

The CAAA of 1977 then required the Prevention of Significant Deterioration (PSD) in NAAQS attainment areas and the New Source Review (NSR) in nonattainment areas. The goal of the PSD provisions was to ensure that the air quality in attainment areas did not degrade. These provisions did not require offsets, but did require sources to employ the BACT for the appropriate source category and class. The goal of the NSR was to ensure that the net emissions from sources in nonattainment areas did not substantially increase; thus, the provisions only allowed the construction or modification of sources if the facilities used the Lowest Achievable Emission Rate (LAER) and offset emissions from other existing sources.[54] The 1977 CAAA also required all major sources

---

[53] The EPA, on a case-by-base basis determines the appropriate emissions rate for a source, or source class or category, that achieves the level of "best available control technology" and they are to take into account energy, environmental and economic impacts, and other costs. See Section 169(3) of the CAA and NRC 2004 and Burtraw and Evans 2004.

[54] The "Lowest Achievable Emission Rate" technology is the most stringent emission limitation possible and the EPA is to determine what this rate is for a source by choosing either (1) the most stringent limitation in any State Implementation Plan or (2) the most stringent limitation achieved in practice for a source in the relevant class and category. See Section 171(3) of the CAA.

in nonattainment areas to install Reasonably Available Control Technology (RACT).[55] These provisions that mainly focused on new and modified sources created a number of problems and distortions, one of which was the extension of the lives of industrial plants and electric generating units beyond their normal lives. For a more detailed discussion of these provisions, their effects, and recent changes to them see NRC 2004, Chapter 5.

The CAAA of 1990 contained a number of provisions that directly limited NO_X and VOC emissions and that empowered the EPA to take further action. Title I of the 1990 CAAA contained provisions that aimed to help all regions of the U.S. attain the ozone NAAQS. These provisions are called the "15% rate of progress plan" and they required that states improve the NO_X and VOC emission rates of their sources in nonattainment areas by 15% from 1990 levels by 1996. States were to achieve this goal by requiring that sources in nonattainment areas use RACT. Title I then mandated a schedule for continued progress that started in 1997. The nonattainment regions must reduce NO_X emissions, VOC emissions, or a combination of both by at least 3% every three years. Some regions have obtained waivers exempting them from this provision, or altering its requirements, when atmospheric modeling showed that NO_X reductions – which were more economically feasible than VOC reductions – would not improve ozone air quality (see Burtraw and Evans 2004 for a more detailed discussion).

The 1990 CAAA also required the EPA to set emission standards for sources that emit more than 10 tons per year of a hazardous air pollutant (HAP). The EPA required Maximum Achievable Control Technology (MACT) standards for Synthetic Organic Chemicals that affected VOC emissions from industrial processes in 1994, 1995, 1997

---

[55] The EPA determines the RACT for source categories and they take cost and other factors into account in this process. They have done this for over 60 such categories (NRC 2004, pg. 186).

and 1999. They implemented various Solvent and Coating Controls in between 1993 and 1996, and in 1998 (U.S. EPA 2005).[56]

The most well known provisions of the 1990 CAAA, Title IV, created the Acid Rain Program. Title IV created a cap-and-trade program to reduce SO_2 emissions from coal-fired electric generating units. It also specified two phases of NO_X emission-rate standards for existing electric generating units. Phase I targeted 265 older coal-fired generating units. Congress intended Phase I to commence in 1995, but litigation delayed its implementation for one year. Phase II commenced in 2000 and required further reductions from Phase I units and reductions from other coal-fired generating units.[57]

The 1990 CAAA also created the Ozone Transport Commission (OTC). Congress recognized that the interstate transport of ozone and its precursors was a problem in the Northeastern and Mid-Atlantic States and created the OTC, a multi-state organization, to help address this.[58] The OTC states agreed that all large stationary sources of NO_X in the region, regardless of ozone NAAQS attainment status, would adopt the "reasonably available control technology" requirements of Title IV of the CAA in 1995. They also adopted a seasonal cap-and-trade program for NO_X from power plants and industrial boilers called the OTC NO_X Budget Program, which commenced in 1999.[59] The program, which the EPA helped implement, capped emissions from affected sources between May 1st and September 30th each year; this period is called the "ozone season" because ozone formation is the biggest problem during this time of year in these states. (Chapter 4 discusses these programs in more detail.)

---

[56] See U.S. EPA 2005 and the EPA's, "Taking Toxics Out of the Air," at http://www.epa.gov/air/toxicair/takingtoxics/p2.html.

[57] Phase I required tangentially-fired coal boilers to reduce their emissions rate below 0.45 lbs/mmBTU and dry bottom wall-fired units to reduce theirs below 0.5 lbs/mmBTU. Phase II required boilers to reduce their NO_X emission rates to between 0.4 and 0.86 lbs/mmBTU, depending on type. It applied to units with capacities greater than 25 MW.

[58] The OTC consists of representatives from Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont, Virginia, and Washington D.C., see Sections 176(A) and 184 of the 1990 CAAA.

[59] All states in the OTC except Maine, Vermont, and Virginia participated.

In light of the updated 1997 ozone NAAQS and the expectation that interstate transport would continue to contribute to ozone nonattainment in Eastern States, the EPA used its authority under the CAA to call for revision of the NO_X State Implementation Plans (SIPs) in Eastern States, in 1998.[60] States could choose to comply with the SIP Call by participating in the SIP Call NO_X Budget Program (NBP), a cap-and-trade program for stationary sources, or by submitting a plan for source-specific NO_X emission rate limits. While the contribution from power plants and other large stationary sources to total anthropogenic NO_X emissions is small relative to that from mobile sources (Section 2.3.1), the EPA focused this regulation on stationary sources for two reasons. The first reason was because the plumes of emissions from high stacks are more prone to interstate transport and the second, because they believed that reductions in NO_X from stationary sources to be the most cost-effective option. For example, in its rulemaking to implement the 1998 NO_X SIP Call, the EPA gave states the option to achieve the required NO_X reductions from any sources (U.S. EPA 1998a, pg. 57378). But the EPA stressed that reductions from stationary source emissions – specifically power plants and other large industrial boilers – through the SIP Call NO_X Budget Program were the most cost effective option compared to reductions from other stationary sources or mobile sources (U.S. EPA 1998a, pg. 57402). All states opted to reduce emissions by participating in the NO_X Budget Program (U.S. EPA 2006b, pg. 3). The program became fully effective May 31st of 2004 after delays from lawsuits. The SIP Call NBP, like the OTC NBP, is a seasonal cap-and-trade program, which caps emissions from May 1st through September 30th from large stationary sources.

---

[60] The EPA's authority to promulgate the SIP Call stemmed from CAA § 110(a)(2)(D). The SIP Call required 22 states and the DC to submit revised SIPs to "prohibit specified amounts of emissions of NO_X – one of the precursors to ozone (smog) pollution – for the purpose of reducing NO_X and ozone transport across State boundaries in the eastern half of the United States," 63 *Federal Register* 57356 (October 27, 1998). The additional participating states, compared to the OTC program, are: AL, IL, IN, KY, MI, NC, OH, SC, TN, VA, WV. Parts of GA and MO will be included in 2007.

In 2005, the EPA used its authority under the CAA §110(a)(2)(D) to issue a broader set of programs that require NO_X and SO_2 reductions in the Eastern U.S. called the Clean Air Interstate Rule (CAIR).[61] This section of the CAA contains a "good neighbor" provision that requires states to ensure, through their SIPs, that sources' emissions do not significantly contribute to the nonattainment of downwind states or to those states' abilities to maintain the NAAQS. The EPA first used this provision to implement the NO_X SIP Call. The EPA's experience with this action, including the fact that the SIP Call withstood substantial litigation, helped them develop and promulgate CAIR.

CAIR calls for SIPs from upwind states whose emissions of the precursors SO_2 and NO_X significantly contribute to the nonattainment of the ozone NAAQS or fine particulate matter ($PM_{2.5}$) NAAQS in downwind states. Upwind states must make specific reductions in SO_2 and NO_X emissions if they contribute to $PM_{2.5}$ nonattainment and must make reductions in summertime NO_X emissions if they contribute to ozone nonattainment. In promulgating CAIR, the EPA used air quality modeling and monitoring to determine the amount of upwind states' emissions that "contribute significantly to downwind nonattainment, or interfere with downwind maintenance" and then required those states to eliminate that quantity of emissions.[62] As, in the NO_X SIP Call, CAIR does not require states to make these reductions in a particular manner or from certain sources. The EPA did, however, develop a model cap-and-trade program that states could adopt if they chose. The rule also encourages states to use an emissions cap approach, even if they do not adopt the model rule, to help give certainty to the amount of reductions over time. Although CAIR will achieve reductions in NO_X of over

---

[61] CAIR applies to 28 eastern states and D.C., see 70 *Federal Register* 25162 (May 12, 2005).
[62] See 70 *Federal Register* 25162 (May 12, 2005).

60 percent from 2003 levels,[63] it is not expected bring all the Northeastern states into full compliance (U.S. EPA 2006b, NESCAUM 2006).

2.6. Summary

This chapter first reviewed ozone formation chemistry and the atmospheric chemistry literature that stresses the importance incorporating the impacts of the timing and location of NO_X emissions into regulations aimed to reduce ozone. Section 2.2 discussed the sources of NO_X, VOC, and ozone emissions. The high concentrations of biogenic VOC emissions in the Eastern U.S. make NO_X reductions important for controlling ozone pollution. Few data are available that describe the geographic and temporal variations in the NO_X from mobile sources. Although mobile sources contribute about 60% of anthropogenic NO_X emissions, the lack of data for mobile source emissions and the monitoring difficulties this reflects makes NO_X from stationary sources an easier target for regulation – especially for differentiated regulation. (Section 6.4.2 discusses how future research on differentiation could address the contribution of mobile source to ozone formation and that this may be worthwhile on the basis of cost-effectiveness).

The trends in ozone concentrations and precursor emissions show that efforts to reduce precursor emission have succeeded and have improved air quality. But there is still a mismatch between the science of ozone formation and the regulation of precursor emissions. The most recent regulations of NO_X in the Eastern U.S. will further reduce the aggregate cap on summertime NO_X emissions from stationary sources. The fact that these reductions are not expected to fully solve the ozone air quality problem provides a motivation to consider whether a cap-and-trade program that differentiated between NO_X

---

[63] CAIR caps annual NO_X emissions from affected sources at 1.5 million tons in 2010 and at 1.3 million tons in 2015. The program caps seasonal NO_X emissions (for the ozone season) at 0.58 and 0.48 million tons in 2010 and 2015 respectively. CAIR will also require reductions in SO_2 of over 70 percent from 2003 levels.

emissions at different times and locations could be implemented, as the literature studying the science of ozone formation recommends.

The Clean Air Act prohibits the EPA from balancing the costs and benefits of reducing concentrations of different secondary pollutions. Differentiation could be a useful regulatory tool within this statutory framework because it could help resolve tradeoffs that occur when reductions in a precursor emission like NO_X cause concentrations of one secondary pollutant (like ozone) to decrease but cause increases in another (like particulate matter). This research focuses on whether differentiated regulations could help areas achieve the ozone air quality standards but it could be extended to assess costs and benefits or applied to other secondary pollutants.

Chapter 3 – Differentiated Regulations in Theory and Practice

The previous chapter discussed two related reasons to consider a differentiated cap-and-trade program for NO_X emissions. First the time and location of NO_X emissions determine their impact on ozone formation. Second, policies that have motivated large reductions in NO_X emissions without specific attention to the timing and location of these reductions have not been able to helping all areas of the Eastern U.S. attain the ozone air quality standards. The environmental economics literature provides another reason to consider differentiated cap-and-trade programs. It suggests that in order to be cost effective, regulations must acknowledge all variables that influence the contributions of emissions to environmental problems. These variables include location, timing, and chemical composition. Theoretically it is possible to design cap-and-trade programs that account for these variables, but history reveals little practical experience with the implementation of regulations that differentiate emissions by characteristics other than their chemical properties. This chapter reviews the environmental economics literature that explains the potential efficiencies associated with using differentiated permit trading programs and the theoretical reasons that these programs can be cost-effective. These concepts could be applied to the case of NO_X emissions and ozone pollution (Section 3.1.1).

This chapter also reviews the challenges to the implementation of differentiated permit trading programs discussed in the literature (3.1.2). These implementation challenges help explain why differentiated regulations have not been used to address the ozone problem. The relevant implementation challenges fall into three categories: 1) technical (modeling and potential emission reductions), 2) economic and technical (transaction costs), and 3) political, legal, and economic (distributional effects). The literature suggests that simplified, second-best, differentiated regulations could address these challenges. However, major permit trading programs have been implemented without differentiation (e.g. the Acid Rain and NO_X Budget Trading Programs). The typical justification for this is that the inefficiencies from the lack of differentiation are small in comparison to the transaction costs from the modeling requirements to support a differentiated program.

Sections 3.2 through 3.2.4 review studies that have evaluated the un-differentiated NO_X cap-and-trade programs. Largely due to simplifications in their estimates of the costs and benefits of these NO_X cap-and-trade programs, these studies offer little guidance on the potential benefits of differentiation. The studies do suggest that although the NO_X cap-and-trade programs have been very successful in achieving targeted NO_X reductions at least cost, the benefits of the regulations would not outweigh their costs were it not for the fact that they caused reductions in particulate matter in addition to some reductions in ozone. This and the discussion in Chapter 2 about tradeoffs between ozone and particulate matter reductions suggests that a better approach is needed to address peak ozone concentrations in the Eastern United States.

Section 3.3 discusses implementation challenges associated with the specific case of a differentiated NO_X cap-and-trade program for stationary sources in the Eastern United States. Chapter 2 suggested that modeling capabilities for predicting ozone concentrations have improved and also that the lack of differentiation may be a detriment to the effectiveness of the cap-and-trade programs now in place. This section first extends the discussion on how advances in air quality forecasting and modeling may now lessen some of the challenges associated with implementing a differentiated cap-and-trade program for NO_X emissions from stationary sources. Technical improvements could also reduce the administrative and transaction costs associated with a differentiated cap-and-trade program, especially if it was applied to electric power plants (Section 3.3.2).

The distributional effects of a differentiated regulation could motivate and/or enable legal challenges that could lessen such a program's effectiveness. Section 3.3.3 uses examples from the legal challenges to past air quality regulations to illustrate the potential challenges associated with the distributional effects of a differentiated cap-and-trade program and the legal mechanisms through which affected parties could challenge and influence it. Finally, this chapter summarizes one conception of a differentiated NO_X cap-and-trade program that could be implemented in the Eastern United States.

3.1. Differentiated permit trading programs in theory

Early work on differentiated permit trading programs focused four topics: 1) the efficiency argument for spatial differentiation, 2) the theoretical ability of spatially differentiated permit trading to produce the least-cost solution to air quality problems 3) the practical difficulties associated with implementation, and 4) potential second-best solutions. The empirical work that this section reviews examined the effectiveness of regulations that utilized second-best solutions and found, generally, that the regulations did not closely approximate the least-cost, differentiated solution.

Dales (1968) and Crocker (1966) independently suggested using emission licenses, or permits, as a means to address environmental externalities.[64] Following this, Montgomery (1972) formalized the concept by proving the existence of equilibriums in markets for two types of emission permits that can theoretically achieve targeted aggregate emission reductions at the least possible cost. One type of permit-trading program created a market for "emissions licenses," where each license conferred the right to emit up to a certain amount. The second type of program created "pollution licenses," which conferred the right to emit until the emissions had a given level of impact on a set of air pollution monitors. In his treatment of these programs Montgomery (1972) showed

---

[64] This is typically attributed to John Dales (1968) but Thomas Crocker (1966) developed similar ideas independently.

that, at least in theory, a permit trading mechanism could deal with the spatial dimensions of pollution problems: a regulatory system that creates a market for each air quality receptor could be a least-cost solution. He also suggested that the market for pollution licenses would be more applicable than that for emissions licenses because the locations and other attributes of emissions sources typically cause their air quality impacts to vary.[65]

Mendelsohn (1986) showed that treating emissions with dissimilar impacts as homogeneous goods causes welfare losses that are directly related to variations in population density and the extent to which the impacts of the emissions differ. Treating dissimilar emissions alike creates a need for costly over-control, regardless of the reasons that the impacts of two unit-quantities of emissions differ (e.g. location, timing, or chemical composition). Even if the impacts of emissions do differ, a unified market for the emissions (i.e. a un-differentiated cap-and-trade program) will equalize marginal abatement costs among sources without regard to the varying impact they have on the environmental problem. Thus too little is spent to control the most harmful emissions and too much to control those that are least harmful. In the un-differentiated case, sources that impact air quality the most in a given target area, like an ozone NAAQS nonattainment area, only reduce emissions if their marginal abatement costs are less than those at all other sources, including those that have little effect on the air quality in a targeted area. The regulator's only control mechanism in this case is the overall cap on emissions; to meet air quality goals in stubborn areas, the regulator must lower the overall cap on emissions until the air quality is acceptable at all locations. This causes two costly problems: over-control at sources emitting emissions with little impact on the area and the necessity of larger aggregate reductions.

---

[65] In fact, he stated, "...the market in pollution licenses will be more widely applicable than the market in emission licenses ... The development of a decentralized system for achieving environmental goals at a number of different locations is the most important contribution of this article," (pg. 396).

Tietenburg (1995) found empirical evidence of these problems by reviewing eight studies that compared the costs of permit trading programs to traditional prescriptive alternatives and a theoretical least-cost program. The evidence from these studies suggested that local situations sometimes cause the costs of un-differentiated permit trading programs to be higher than traditional prescriptive regulation, in addition to the least-cost theoretical program. He separated the costs of different regulatory options into their "equal-marginal-cost" and "degree-of-required-control" components for each study (pg. 99). While permit-trading programs prevailed over prescriptive regulations in their ability to reduce costs by equalizing marginal control costs across sources, he found that they sometimes required a higher degree of control to achieve the same improvements in air quality. In some cases, depending on the importance of spatial considerations, the un-differentiated permit-trading programs required larger aggregate reductions to such an extent that the additional costs outweighed the cost-savings from equalizing marginal control costs across sources.

3.1.1. Applicability of theory to NO_X emissions and ozone formation

Nitrogen oxide emissions, in their relationship with ozone formation, fit the characteristics of emissions that should be treated as heterogeneous goods. Mendelsohn (1986) suggests that the primary factors that cause one unit-quantity of emissions to differ from another include chemical composition and the time and location of their release. In the case of NO_X, the latter two forms of differentiation are the most important. Although the category of nitrogen oxide emissions does include different compounds, like NO and NO_2, both stationary and mobile sources predominantly emit NO emissions and these are converted to NO_2 after release (Army Corp 1988). The dependence of ozone formation on meteorology makes the timing of emissions important. NO_X emissions at night in cool, calm weather may cause little ozone formation either locally or downwind; but, NO_X emissions on a hot, sunny day can cause a large amount of local ozone formation, as can NO_X emitted either at night or during the day in windy conditions carries downwind to areas conducive to ozone formation. Meteorology or atmospheric chemistry may mean that distant sources contribute most to ozone formation in densely populated areas, but research shows that some sources of NO_X have a larger impact than others (e.g. Mauzerall *et al.* 2005 and Tong *et al.* 2006).

Very recent experience with the seasonal NO_X cap-and-trade programs in the Eastern U.S. suggests that the locational and temporal impacts of NO_X emissions are beginning to diminish the cost-effectiveness of the largely un-differentiated cap-and-trade programs. The OTC and SIP NO_X Budget Programs have successfully reduced aggregate emissions and helped improve air quality in many areas of the Eastern United States (Section 2.4). The 31 highly populated non-attainment areas that remain, however, are driving regulators to further reduce the aggregate emission caps, with the Clean Air Interstate Rule (CAIR), for example, even though modeling predicts that the lower caps will not ensure attainment of the ozone NAAQS in all locations (U.S. EPA 2006b, NESCAUM 2006).[66] In this case, the un-differentiated cap-and-trade programs have worked well to reduce ozone concentrations generally across the Eastern United States. Theory warns, however, of the potential for over control in many areas if regulators address the remaining nonattainment areas with further reductions in the aggregate cap.

3.1.2. Overview of implementation challenges from the literature

Despite recognition of the potential efficiencies associated with differentiated regulations, practical challenges have limited the feasibility of their implementation. The early literature focused on the technical and economics barrier associated with the high costs and uncertainties involved in modeling source-receptor relationships and how they

---

[66] Although CAIR also aims to reduce particulate matter pollution, to which NO_X emissions contribute, it contains provisions that directly address ozone nonattainment by mandating summertime caps on NO_X emissions from stationary sources. It is the actions taken specifically to reduce ozone, which are above and beyond those aiming to reduce PM, with which we are concerned.

change in time, which are especially challenging for systems of many sources and receptors (e.g. Atkinson and Tietenberg 1982, Krupnick *et al.* 1983, Mendelson 1986, Tietenberg 1995). Section 3.3.1 discusses how advances in weather and air quality forecasting may now mitigate these challenges.

There are also political and legal barriers to the implementation of differentiated regulations. Political and legal disputes are likely to arise over the definitions of boundaries and rules for permit exchange in a differentiated program. Industry and state governments should be expected to dispute these definitions because small changes could yield large benefits (e.g. Mendelson 1986, pg. 309). Regulated industries could base these disputes on claims of fairness (e.g. Atkinson and Tietenberg 1982, Mendelson 1986). Industries and state governments could also contest boundary definitions using arguments over uncertainties in air quality models. Section 3.3.3 reviews experience with the implementation of recent cap-and-trade programs, which suggests the latter as a likely outcome.

Another technical implementation challenge is whether sources can reduce emissions in response to incentives that vary in time and by location. For example, could incentives change driving patterns in cities? Or could the operators of power plants reduce emissions in the short run by changing which power plants supply power in particular hours and locations? Chapter 5 presents estimates of how much flexibility power plants have to provide short-term NO_X emission reductions during periods when ozone formation is likely to be a problem.

3.1.3. Second-best alternatives to full differentiation

In the face of these practical challenges, there are second-best alternatives to full differentiation. Desirable qualities for second-best regulations include: costs that approach the least-cost theoretical program, cost-effectiveness independent of the initial allocation of permits, minimal information needs for regulators, low transaction costs, and no unnecessary restrictions on the trade of permits. Air quality should also improve to meet the standards in areas where it is poor and it should not significantly degrade in other areas. A number of papers have proposed systems with the theoretical potential to meet these requirements. Regulators have implemented programs with some of these characteristics, but have most frequently implemented un-differentiated permit trading programs, themselves a second-best alternative.

For problems in which spatial differentiation is a concern, one type of second-best alternative limits the geographic size of permit markets and another type limits the transactions allowed within a broader market. Programs that limit the geographic size of markets might use air quality modeling to define a number of "zones" in which emissions would be treated alike and exchange rates at which sources in different zones could trade permits. If regulators had sufficient information to correctly define these zones and exchange rates, their use could theoretically improve cost-effectiveness compared un-differentiated regulations (Tietenburg 1995). In addition, significant increases in cost-effectiveness could come from the designation of just a few zones. For example, Roach *et al.* 1981 found that airshed-level zones for SO_2 permit trading in the Southeastern U.S. reduced the cost penalty from over control by a factor of three to four compared to a state or regional program. McGarland (1984) found that dividing the airshed in the area of Baltimore, MD into three zones for PM pollution cut the over-control cost penalty in half (cited in Tietenburg 1995).

A potential problem with a zonal system is that small zones limit trading opportunities, especially if trading between zones is not allowed; this could increase costs and create potential for market power (Atkinson and Tietenberg 1982, Krupnick *et al.* 1983). These are not necessarily problems if regulators can correctly designate the zones. Tietenburg (1995), citing ICF (1989), noted that even if regulators must define small zones to achieve local air quality, the option of trading may still be desirable because even trading between generating units in a single power plant can generate cost-savings compared to the case of no trading.[67] In the realistic limited-information case, however, when regulators cannot be expected to correctly designate zones or when the dispersion characteristics of sources within the zones vary, the zonal approach creates the potential for unnecessarily limiting trading and thus for increased costs (Atkinson and Tietenberg 1982, Krupnick *et al.* 1983, McGartland 1984).

Another problem with the zonal approach is that some of its administrative practicality comes from the assumption that regulators could define stable, predefined zones. In fact, because populations, activity patterns, sources, and meteorology change in time, so will the proper zone definitions and exchange rates. In addition, in order for the regulators to efficiently allocate permits to the zones initially, they must know the abatement-cost characteristics of the sources in each zone and how they change over time (Atkinson and Tietenberg 1982). This system places a large burden on the regulators to correctly determine, and update, the initial allocation of permits to the zones (Atkinson and Tietenberg 1982, Krupnick *et al.* 1983).

Rather than narrowing or dividing the market to deal with locational problems, another second-best option is to limit transactions within a broadly defined market. The literature has proposed three types of programs: the nondegradation offset (Atkinson and Tietenberg 1982), the pollution offset (Krupnick *et al.* 1983), and the modified pollution offset (McGartland and Oates 1985). The authors developed these programs because the zonal approach places a high burden on regulators to correctly determine boundaries and exchange rates, and because of transaction cost and other problems with Montgomery's systems of pollution and emission licenses.

---

[67] Also see Ellerman (2000) for a discussion of autarkic compliance with Phase I of the Acid Rain Program cap-and-trade for SO_2. While not the most efficient option compared to wider, market-based trading, within-utility trading did generate cost saving.

Montgomery's system of pollution licenses creates high transaction costs for industry because it requires firms to hold and trade permits in a market for each receptor that their emissions affect (Atkinson and Tietenberg 1982, Krupnick *et al.* 1983). The system of pollution licenses does have the politically helpful attribute that the initial allocation of pollution permits to sources does not affect the outcome; trading in an efficient market could achieve the least-cost attainment of air quality goals. The flexibility in initial allocation is an important characteristic for permit-trading programs because it increases political feasibility (e.g. Krupnick *et al.* 1983).

Montgomery's alternative system of emission licenses reverses these problems: it has lower transaction costs because sources could simply trade emission permits in a single market, but it requires regulators to determine an initial allocation of permits with certain characteristics. In his emission licenses system, Montgomery requires that trades do not result in diminished air quality at any receptor (a "nondegradation condition") so that its air quality outcome corresponds to that of the system of pollution licenses. For the equilibrium in this market to both satisfy the nondegradation condition and achieve air quality goals at least cost, the regulators are required to determine an initial allocation of permits such that if all permits were used, the air quality would just bind at each receptor.[68] There are two reasons for this. First, if the allocation allowed too much pollution at a receptor, the program might not achieve the air quality goals. Second, if it restricted emissions too much at a receptor the "nondegradation condition" would not allow the emissions to increase and the program would result in over control, not the least-cost solution (Krupnick *et al.* 1983). In order for the regulator to determine the proper initial allocation, they would need to know the source's abatement-cost characteristics and the source-receptor transfer coefficients and use them to solve for the

---

[68] For further discussion of the "nondegradation condition" and the burden it would place on regulators see McGartland and Oates 1985 and Krupnick *et al.* 1983, who also discuss why this condition is overly stringent.

efficient level of emissions at each source. This is clearly not feasible, and if it were, regulators would no longer need a market-based system.

Krupnick *et al.* (1983) and McGartland and Oates (1985) show, however, that Montgomery's "nondegradation condition" is overly restrictive. Krupnick *et al.* proposed a system in which sources trade emission permits as in Montgomery's system, but with the requirement that air quality not exceed a desirable standard at each receptor (which might vary by receptor in accordance with population density). This, in practice, could simply mean that a new source or a source wanting to increase emissions would have to purchase offsets for the added emissions from other sources such that the air quality at the relevant receptors did not degrade. They call this a system of pollution-offsets and show that its ability to achieve the air quality goals at least-cost does not depend on the initial allocation of permits.

Environmentalists criticized the system envisioned by Krupnick *et al.* (1983) because it could lead to the degradation of air quality at locations with initial air quality below the standard. McGartland and Oates (1985) proposed a system of pollution-offsets that slightly altered the one conceived by Krupnick *et al.* (1983). They strengthened the requirement that air quality not exceed the standard at any location by requiring that air quality neither exceed the standard *nor* degrade at any receptor. In other words, they replaced the air quality standard at attainment receptors with a new standard equal to the current level of air quality. They showed that this system could achieve the least-cost solution independent of the initial allocation of emission permits.

The CAA applied this concept using "bubble" provisions that allowed sources to forego required emission reductions if they secured equal reductions at a nearby source (e.g. Tientenburg 1990). The "offset" provisions allowed new sources to enter a nonattainment area only if they secured more than enough offsetting emission reductions from existing sources in that area, as this was believed to guarantee improved air quality in the nonattainment area.[69]

The final, most frequently used, second-best alternative is to ignore the need for differentiation entirely. Regulators have most often chosen this option, for example with the Acid Rain Program and the NO_X Budget Programs. Because policy makers have chosen to implement un-differentiated programs, the more recent literature has focused on determining whether the lack of differentiation has lessened the effectiveness of the currently operating cap-and-trade programs.

3.2. *Effectiveness of un-differentiated NO_X programs*

Modeling requirements have made it difficult for regulators to implement even second-best, differentiated regulations. The major federal cap-and-trade regulations in the U.S. – the Acid Rain Program, SIP Call NO_X Budget Program, and CAIR – all but ignore the spatial, and temporal, aspects of the air quality problems they target. The SIP Call NO_X Budget Program, like the OTC program before it, does differentiate between the "ozone season" when ozone formation is a problem and the rest of the year. This section reviews analyses of the un-differentiated cap-and-trade programs that have been implemented to reduce NO_X emissions. Some studies have estimated the costs and benefits of these regulations. However, few studies address whether differentiation could improve the performance of these regulations and the simplifications of the studies that do mean they offer little guidance on the potential costs of the un-differentiated nature of the regulations or on the benefits of differentiation.

---

[69] See 44 *Federal Register* 71780 (11 December 1979) and 40 CFR 51 Appendix S, originally presented in 44 *Federal Register* 3274 (I6 January 1979) and also see 51 *Federal Register* 43814, "Emissions Trading Policy Statement."

3.2.1. Geographical "hotspots" and implications for differentiation

While the earlier literature focused on potential efficiency gains from addressing the locations of emissions, more recent evaluations of the trading programs that have been implemented ask whether the lack of spatial differentiation has lead to problems such as "hotspots" (e.g. Swift 2004, Burtraw *et al.* 2005). Hotspots in an emissions trading program are areas (or times) in which the majority of sources bought permits to cover their emissions, rather than made reductions. Swift (2004) found that hotspots did not occur in the Acid Rain Program's SO_2 cap-and-trade program. Reviews of the OTC and SIP Call NO_X Budget Trading Programs found "very little" state-level shifting of emissions to concentrated areas (Swift 2004 and Farrell 2003). Farrell (2003) found that the OTC NO_X Budget lowered average and peak emissions in equal proportions between 1998 and 2000. An EPA and OTC analysis shows that under the OTC NO_X Budget both daily total emissions and daily peak emissions declined since 1997 (OTC 2003 pg. 8). Swift (2004) found, in addition, that the largest NO_X emission sources abated the most under the OTC program.[70]

These studies do not, however, directly address the complexity that the lack of NO_X emission hotspots, especially at a state-level of aggregation, does not mean that all the NO_X reductions were made at the most effective locations or times. For ozone, large emission reductions from the largest sources are not necessarily the most damaging in terms of ozone formation and transport in emission plumes from power plants (Ryerson 2001, Mauzerall *et al.* 2005). Ozone formation chemistry is sufficiently complicated that, despite the absence of emissions hot spots and the programs' effectiveness in many less-populated areas of the Eastern U.S. thus far, the lack of differentiation in seasonal NO_X

---

[70] Ellerman (2004) offers an explanation of similar findings for the Acid Rain Program: the largest sources are the least expensive sources of abatement, on a per-ton basis, when there are large capital costs associated with installing emission control technologies (pg. 86).

cap-and-trade programs limits their potential to completely solve the ozone air quality problem in all locations.

Research rooted more in the field of atmospheric chemistry than in economics has begun to examine the effects of particular NO_X sources on the formation of ozone and the potential benefits of reducing NO_X at these sources and others similar to them. For example, Mauzerall *et al.* (2005) examined the health impacts of emissions from particular point sources. They found that the time and location of NO_X emissions caused the amount of ozone formation to differ by up to a factor of five. This suggests that locational and temporal details are extremely important for ozone. These studies suggest that NO_X cap-and-trade programs in the Eastern U.S. fit the predictions of the early theoretical literature: un-differentiated regulations may be leading to significant over-compliance costs at sources distant from the remaining areas of poor air quality while not requiring reductions at the sources contributing most to nonattainment.

The transaction costs from differentiation and the excess abatement costs from a lack of it create a tradeoff (Krupnick *et al.* 1983). As the problem's spatial (or differentiating qualities) become more important, the excess abatement costs from not differentiating can overwhelm the transaction costs. This may be the case for both PM and NO_X, and the transaction costs of a differentiated program may be decreasing because of improvements in modeling capabilities (Krupnick *et al.* 1983, Mauzerall *et al.* 2005, Tong *et al.* 2006).

The modeling requirements extend beyond the analysis necessary to implement a differentiated regulation. Detailed modeling is also required in order to determine the benefits of NO_X reductions made under alternative reguations. It is difficult to say that the inefficiencies from the lack of differentiation are small based on studies that may overestimate the benefits of nondifferentiated cap-and-trade programs because they do not accurately model the effects of the nondifferentiated NO_X reductions on air quality.

The next two subsections examine how simplifying assumptions may mask inefficiencies caused by the lack of differentiation. Furthermore, none of the benefit and cost analyses of NO_X regulations to reduce ozone in the Eastern U.S. account for the actual compliance behaviors taken by sources in response to the regulations because none were *ex post* analyses.

3.2.2. Costs and benefits of the OTC NO_X Budget Trading Program[71]

There are no studies that have performed *ex post* estimates of the costs or benefits of the OTC NO_X Budget Program. Farrell *et al.* (1999) estimated the costs of the OTC NO_X Budget cap-and-trade program before it was implemented and found the cap-and-trade approach would be more cost-effective than an alternative command-and-control regulation.[72] They predicted an annual average cost of $161 million (2000$) for the cap-and-trade system and of $302 million for the traditional command-and-control approach.

Although Farrell *et al.* (1999) modeled a range of control technologies including post-combustion and combustion controls, their model did not account for possible changes in the utilization or operation of power plants (pg. 114). Historical evidence shows that power plants have used small combustion controls and changes in utilization in response to the price on NO_X emissions created by the OTC cap-and-trade program. This omission likely caused the Farrell *et al.* estimates of the costs of the program to be slightly high.[73] They also predicted marginal control costs under the OTC NO_X Budget Program to be $1461/ton in 1999 and $1887/ton in 2002 (2000$). The price of NO_X allowances fell below $1000/ton by the end of 1999 after the market settled and remained

---

[71] All dollar values in this section and the next two have been adjusted for inflation using the consumer price index (CPI) and are reported in year 2000 dollars.

[72] Farrell *et al.* (1999) also note that their models did not consider the temporal effects of NO_X on ozone formation and whether or not a NO_X cap-and-trade program would address the episodic nature of ozone formation (pg. 122).

[73] Industry publications that expected power plant dispatch to be a short-term compliance option (such as C. Seiple and R. LaCount, "NO_X Emissions Trading: Changing Generator Behavior?" *Public Utilities Fortnightly*, July 15, 1999.)

between about \$500 and \$1700/ton – but mostly below \$1000/ton – through 2002, again suggesting that the Farrell *et al.* cost estimates were slightly high.[74]

3.2.3. The costs and benefits of the SIP Call NO_X Budget Program

Three studies have estimated the projected costs of the SIP NO_X Budget Program (U.S. EPA 1998b, Burtraw *et al.* 2001, and Krupnick *et al.* 2000). Krupnick *et al.* (2000) studied both the costs and benefits of reducing NO_X emissions for a 12-state region that they felt represented the 22-state region the EPA was considering for the SIP Call NO_X Budget Program. Their benefit-cost analysis suggested that the EPA's target reductions for the NO_X SIP Call were reasonably close to the optimal reductions, if the mortality risks from both ozone and particulate matter were included in the calculation of the benefits from reducing NO_X emissions (pg. 24).

Krupnick *et al.* (2000) also compared the costs of controlling emissions with a cap-and-trade program to a command-and-control approach and to a spatially differentiated cap-and-trade program based on ozone exposures. They found that the costs associated with a cap-and-trade program would be about 50% lower than under a command-and-control scenario that required a similar level of abatement.[75] They found that, to achieve the same reduction in population-weighted ozone concentrations, a program in which sources traded population-weighted ozone exposures instead of NO_X emissions would have slightly lower costs than the un-differentiated NO_X emissions trading program.

Although some have taken the Krupnick *et al.* (2000) results as evidence that a more spatially differentiated cap-and-trade program for NO_X would not create efficiency

---

[74] For a more complete discussion of NO_X allowances prices, see, for example, Environmental Finance's "Success in the US" at http://www.environmental-finance.com/2004/0410oct/emission.htm and (OTC 2003).

[75] The command-and-control approach that they modeled required electric utilities to reduce NO_X emissions to either 0.15 lbs/mmBTU or an 85% reduction, whichever was lower; other sources were required to meet the lower of 0.15 lbs/mmBTU or a 70% reduction.

gains (Burtraw *et al.* 2005), omission of details about compliance possibilities and source locations in Krupnick *et al.* (2000) make this inconclusive, especially given the complicated chemistry of ozone formation. The models that Krupnick *et al.* (2000) employed only crudely linked the emissions of NO_X at particular locations to the formation of ozone at others. They developed source-receptor coefficients[76] for a small number of regions that were separated by wind patterns and large geographical formations (like the Appalachian Mountains), but that largely followed state lines (pg. 7). They did not consider control strategies like fuel switching and decreased utilization and note that the inclusion of these options may have further reduced the costs of the ozone-exposure trading program compared to the NO_X emissions trading program (pg. 24).

Burtraw *et al.* (2001) and U.S. EPA (1998b and 1998c) estimated *ex ante* costs and benefits for the SIP Call NO_X Budget Program. Burtraw *et al.* (2001) used methods very similar to those used by Krupnick *et al.* (2000) and therefore are also subject to the criticism that they did not model details of actual sources and their locations. Burtraw *et al.* (2001) considered a 19-state (plus DC) region similar to the EPA's 22-state (plus DC) NO_X SIP Call region, but did not explicitly estimate the ozone-related health benefits associated with reductions NO_X emissions.

Despite differences in some assumptions, the three studies' estimates of average costs and benefits were reasonably close for reductions in the ozone season (in 2000\$). The EPA estimated average costs of NO_X reductions under the program to be \$1,984 per ton; Krupnick *et al.* (2000) found average costs of \$1,987 per ton; and Burtraw *et al.* (2001) found average costs of \$2,321 per ton.[77] The estimates in Burtraw *et al.* (2001) are

---

[76] Their source-receptor coefficients describe how much NO_X emissions from one area contribute to population-weighted ozone formation in another (Krupnick *et al.* 2000, pg. 7).

[77] The EPA did not model marginal costs. Burtraw *et al.* (2001) and Krupnick *et al.* (2000) found these to be \$3,649 per ton and \$4,250 per ton respectively (2000\$). The Krupnick *et al.* marginal costs are likely higher because they did not model compliance options like changes in output and changes in operations, although neither study included the possible use of combustion controls like low-NO_X-burners.

higher because they do not model the possibility of combustion controls and also model a slightly larger reduction in emissions.

Burtraw *et al.* (2001) also found that extending the cap-and-trade system annually, instead of just for the summer months, would increase its net benefits and cost-effectiveness. The benefits associated with reductions in PM created the benefits outside the ozone season, as well as substantial benefits within the ozone season. Similarly to Krupnick *et al.*, Burtraw *et al.* (2001) found that neither the PM nor the ozone-related benefits alone caused the net benefits from the modeled seasonal NO_X cap-and-trade programs in the SIP Call region to be positive, although the combination did.

Burtraw *et al.* (2003) analyzed 18 scenarios and found that, in all cases, the net benefits of an annual cap-and-trade program to be at least those from a seasonal program because NO_X reductions moderated particulate matter concentrations outside the ozone season. They evaluated 18 scenarios in order to address the large uncertainties surrounding some of the model parameter values in previous studies (including Burtraw *et al.* 2001). In their scenarios, they address three issues: assumptions about future deregulation in electricity markets, epidemiological uncertainty regarding premature mortality, and uncertainties surrounding economic valuation of mortality risk (Burtraw *et al.* 2003, pg. 383). They again, however, neither modeled the possibility that power plants could use combustion controls to reduce NO_X nor the ozone-related health benefits that could result from NO_X reductions. They found that the omitted benefits, under midpoint assumptions, would need to be at least \$1,288 per ton of NO_X reduced in the SIP Call region seasonal cap-and-trade scenario for the policy's net benefits to be positive (2000\$, pg 397). Mauzerall *et al.* (2005) estimated mortality impacts of between \$12,090 and \$59,660 per ton of NO_X emitted, depending on the location and timing of the NO_X emissions and their subsequent impact on ozone formation. They estimated the morbidity effects to be between \$72 and \$356 per ton (2000\$, pgs. 2861-3).

Although Burtraw *et al.* (2001 and 2003) focused only on their point that the annual cap-and-trade program for NO_X has higher net benefits than the seasonal regulation, another important point can be derived from their analysis. If the seasonal cap-and-trade program does not provide positive net benefits from the reduction of PM alone, then the seasonal cap-and-trade program must reduce ozone concentrations for its net benefits to be positive. That is, if the seasonal cap-and-trade program is not effective because it does not account for the time and locational variation in the impact of NO_X on ozone formation, the ancillary benefits from PM reduction are not enough to make it efficient – so it is important that the program actually achieve its intended ozone concentration reductions. It should be noted again, however, that Burtraw *et al.* (2001 and 2003) may overestimate the costs of the program because the do not consider combustion controls as an option that sources can use to reduce NO_X emissions.

3.2.4. Efficient levels of NO_X emissions

Banzhaf *et al.* (2004) modeled both marginal benefit and cost curves for the reduction of NO_X and SO_2 from power plants in the U.S. in order to estimate efficient levels of these pollutants, under a cost-effective regulation like a tax or cap-and-trade program in the United States. The Banzhaf *et al.* (2004) study is important because few studies have tried to estimate the most efficient level of emission caps, although policymakers have often set these caps. They found the efficient level of a national, annual cap on NO_X emissions was about 1.4 million tons (between 1.0 and 2.8 million tons) in 2010. The EPA expects that national, annual NO_X emissions after the implementation of the Clean Air Interstate Rule (CAIR) will be about 2.4 million tons.[78]

Three problems with this study are that 1) it did not consider the benefits of reducing ozone through reduced NO_X emissions, 2) it did not incorporate detailed

---

[78] U.S. EPA, "Projected Annual NO_X Emissions from Power Plants with the Final Clean Air Interstate Rule," March 2005, at http://www.epa.gov/interstateairquality/charts_files/cair_emissions_costs.pdf.

information on the locations and emissions characteristic of actual power plants,[79] and 3) it did not consider combustion controls like low-NO_X-burners as an option for the control of NO_X emissions (pg. 7). This study only considered the damages that NO_X emissions caused through NO_2 and the formation of particulate matter; they did not consider ozone-related damages. Their models only allowed plants to install post-combustion controls (selective catalytic reduction and selective non-catalytic reduction) or to decrease utilization in response to power and emission prices. The combustion controls are a lower-cost option for compliance than the post-combustion controls; hence, the inclusion of the former would probably have increased the amount of NO_X abatement that they would have found as efficient.[80] Due to the omission of important details, this study and those discussed in the previous two sections offer little guidance on whether differentiation for NO_X emissions may be worthwhile. For this reason, persistent non-attainment of the ozone air quality standards and potential efficiency benefits are still the primary motivation for considering a differentiated regulation.

3.3. *Differentiated permit trading programs in practice*

Although it is theoretically possible and potentially worthwhile to design a differentiated permit trading mechanism that addresses the time and locational impacts of emissions, there are technical, organizational and political challenges associated with implementation. Challenges that have limited the use of highly differentiated permit trading programs are: 1) technical challenges associated with atmospheric chemistry modeling; 2) technical challenges associated with the potential response of emission

---

[79] They used an electricity sector equilibrium model (Haiku) to create a representative power plant for each of the 13 NERC subregions in the United States. They then aggregated the emission data from these representative plants to the state level and input them into the Tracking and Analysis Framework (TAF). TAF then estimated the resulting pollutant transport, deposition, the formation of secondary particulate matter (but not ozone formation), the resulting human health effects and their monetary value at the state level (Banzhaf *et al.* 2004, pgs. 6-9).

[80] The addition of combustion controls in the power sector model would have altered the marginal cost curve that they estimated, presumably making it increase less steeply and thus intersect with the marginal benefits curve at a higher level of abatement (see Banzhaf *et al.* 2004, pg. 23 for a figure of their marginal benefit and cost curves).

sources to differentiated incentives; 3) the administrative and transaction costs of implementation; and 4) political and legal arguments about fairness and scientific uncertainty motivated by distributional affects.

This section discusses the first, third, and forth of these challenges and leaves more detailed analysis of the second to Chapter 5. Sections 3.2.1 and 3.2.2 discuss how advances in air quality modeling and forecasting and experience with cap-and-trade programs may now mitigate the first and third of these problems. The analysis presented in Chapter 5 suggests that it is possible for power plants to respond to differentiated incentives with short-term emission reductions. Section 3.2.3 discusses how the responses of industry and state governments to the 1998 NO_X SIP Call and to CAIR suggest that legal arguments motivated by distributional effects would occur during the implementation of a differentiated cap-and-trade program. These disputes could constrain the effectiveness of the differentiated regulation if they altered the definitions of when and where NO_X reductions were needed.

3.3.1. Weather and Air Quality Forecasting

Developments in weather and air quality forecasting have lessened the challenge of determining the impacts of individual sources on the air quality problems in nonattainment areas. State-of-the-science air quality models like the EPA-contracted Community Modeling and Analysis Quality (CMAQ) modeling system have achieved major improvements compared to the urban airshed models used for policy in the 1970s and the regional models using in the 1980s (like the Regional Acid Deposition Model used in the preparation of the 1990 CAAA). A “multiscale” model incorporates a mesoscale model of meteorology and a detailed model of emissions from individual sources including the unique characteristics of each plume. This enables detailed analysis of the transport, chemical transformation, and deposition of pollutants and the reproduction of historical conditions, including ozone episodes at detailed scales both spatially and temporally (e.g. O'Neill *et al.* 2006, Tong *et al.* 2006, Eder and Yu 2006).

The EPA and NOAA have also developed the National Air Quality Forecasting Capability. This system combines NOAA's weather forecasting ability and the air quality modeling capabilities like CMAQ to forecast hourly air quality so that people can take action to limit their exposure to poor air quality (e.g. Davidson *et al.* 2004). An evaluation of the air quality forecasting system for New England revealed that the forecasts preformed well, but with room for improvement (Kang *et al.* 2005). The study evaluated three different models that incorporated weather forecasts and atmospheric chemistry information. They found that each overpredicted ozone concentrations. Each model did achieve accuracy of greater than 90% when predicting whether or not the ozone concentration in an hour would exceed or not exceed the 1-hour ozone standard, and between 76 and 90% for the 8-hour standard. However, the general accuracy metric does not give a complete picture of the potential role of the forecasting in air quality policy. For policy purposes, it is important to predict infrequent exeedances of ozone air quality standards and the large number of nonexceedance-hours influences the general accuracy metric. The models correctly forecast between 6 and 36% of exeedances and correctly forecast 64 to 87% of the nonexceedances (pgs. 1790-91).[81]

3.3.2. Administrative and transaction costs

A number of technical improvements could now make the administrative and transaction costs of a differentiated cap-and-trade program tractable. Examples are the continuous emissions monitoring systems that track emissions from stationary sources, and electronic toll systems that could be used to create differentiate incentives for driving in time and space. Experience with cap-and-trade programs and with bid-based dispatch

---

[81] Researchers at the Center for Energy and Environmental Policy Research at MIT are currently performing further research into the ability of air quality modeling and weather forecasting to correctly predict ozone episodes in the Eastern U.S., this research is out of the scope of this dissertation.

and day-ahead and real-time markets for wholesale electricity could also help enable a differentiated cap-and-trade program. The operators of power plants, in particular, use sophisticated weather forecasting and information on fuel prices that can fluctuate in time to make daily decisions to bid generation into power pools. An exchange rate for emissions permits that was announced 48 hours in advance would not greatly change this process, but would rather be another addition factor for consideration in formulating bids.

3.3.3. Legal disputes over distributional effects

Recent experience with CAIR suggests that the third problem – the legal fights over distributional effects – deserves attention. In their legal basis for CAIR the EPA was required to use air quality modeling to determine that certain areas' sources contributed significantly to the nonattainment of the ozone (and PM) NAAQS in particular areas of the Eastern United States. The literature predicts such a process could be contentious (e.g. Atkinson and Tietenberg 1982, Mendelson 1986, Joskow and Schmalensee 1998).[82] States had a stake in the determination of "significant contribution" because the EPA required states to make reductions of the amount that their emissions contributed significantly to downwind non-attainment or "interfer[ed] with downwind maintenance."[83]

Indeed, with CAIR, the process of determining whether states' sources contributed significantly to other state's air quality problems was contentious and the disputes focused around boundaries and the modeling that supported them.[84] The disputes

---

[82] Joskow and Schmalensee (1998) examined the political economy of the allocation of allowances for the Acid Rain Program's SO_2 trading provisions and found the process to be both contentious and complicated. They summarized the underlying issue nicely: "Because emissions permits are valuable and decisions about their distribution are made by political institutions, these decisions are likely to be highly politicized, reflecting rent seeking behavior and interest group politics. ... In particular, little attention has been devoted to how interest group politics and associated rent-seeking behavior affect the allocation of permits in a tradable permit system. This is a serious gap in the literature. The political acceptability of market-based mechanisms for internalizing environmental externalities will depend heavily on their distributional implications," (pg. 4).

[83] EPA, "Rule to Reduce Interstate Transport of Fine Particular Matter and Ozone (Clean Air Interstate Rule)," 70 *Federal Register* 91, pg. 25162.

[84] See the petitions the EPA received to reconsider their findings at http://www.epa.gov/cleanairinterstaterule/rule.html.

caused the EPA to feel obliged to supplement their initial modeling that used the "Regional Model for Simulating Aerosols and Deposition" (REMSAD).[85] Although the EPA stated that it did not agree with criticisms about the REMSAD modeling, it used another model as a response, the CMAQ discussed in Section 3.3.1. This model is publicly available, peer-reviewed, and considered "state-of-the-science", which helped lend credibility to the EPA's designations of significant contribution.

In the case of a differentiated cap-and-trade program, small changes in the definitions of trading ratios or boundaries could have large impacts on particular companies or states. Even with the improvements in modeling, uncertainties still persist. As occurred in the case of CAIR, companies would dispute the EPA's modeling and the resulting definitions of trading ratios by offering their own modeling results that showed a lesser responsibility for the air quality problems. Legal disputes could then delay the program or could lead to changes in the definitions of the program that could hinder its effectiveness. Similar disputes over distributional affects occurred with the allocation of emission permits for the Acid Rain Program and the NO_X Budget Trading Programs (Joskow and Schmalensee 1998). In these cases the resulting allocation has had little, if any, impact on effectiveness of these programs since sources have met the emission caps at low cost. But, the definitions of exchange ratios and boundaries could impact the success of a differentiated regulation. The concept relies on such definitions to more accurately address the environmental problem, beyond just reducing emissions to below and annual or seasonal cap.

---

[85] It is a photochemical grid model that uses atmospheric specie mass continuity equations. Details on this model are available in "Notice of Proposed Rulemaking Air Quality Modeling Technical Support Document", see EPA, "Rule to Reduce Interstate Transport of Fine Particular Matter and Ozone (Clean Air Interstate Rule)," 70 *Federal Register* 91, pg. 25162.

3.3.4. A time- and location-differentiated cap-and-trade program for NO_X

This dissertation contributes to a broader research effort that is seeking to determine whether it would be more cost-effective to directly address the temporal and locational impacts of NO_X emissions from power plants rather than to further reduce the aggregate cap.[86] Weather forecasting models would initiate the regulatory system; they would provide advance warning of the times the formation of high ozone concentrations in critical receptor areas (for instance, those in non-attainment) was likely. The models would also predict the locations of the precursor NO_X emissions with the largest impact on ozone formation during the critical times at the critical receptor areas. Power plant operators would then be notified of the times and locations when a pre-set allowance surrender ratio greater than one-to-one would be imposed on NO_X emissions. Generators could then modify their bids in the day-ahead and real-time wholesale power markets in response to the higher cost of NO_X emissions. The day-ahead and real time markets would then lead to patterns of locational prices that reflected the prevailing NO_X emissions permit exchange rates and result in an altered generator dispatch, compared to that without NO_X prices, and NO_X abatement.

The effectiveness of the system rests on four necessary conditions. The first is that weather and atmospheric chemistry forecasting can predict the conditions conducive to ozone formation with sufficient accuracy and lead-time (at least 48 hours) to influence electricity markets. The second is that the spatial zones and time intervals in which the surrender ratio for the NO_X emissions permits would be varied can be identified with sufficient regularity that a reasonably simple and stable system of differentiated permit exchange rates triggered by reliable and transparent indicators of weather and atmospheric chemistry can be implemented. The third is that there exists sufficient

---

[86] The research is a project of the Center for Energy and Environmental Policy Research at MIT. See also: Martin, K.C., P.L. Joskow, and A.D. Ellerman (2007), "Time and Location Differentiated NO_X Control in Competitive Electricity Markets Using Cap-and-Trade Mechanisms," *Center for Energy and Environmental Policy Research Working Paper.*

flexibility in the redispatch of generating units of differing NO_X emissions rates and in NO_X emissions control that significant NO_X reductions can be accomplished on relatively short notice and without violating transmission network and supply/demand balance constraints. The fourth condition is then that the magnitudes of NO_X reductions that could be effected in the specified areas and times in response to differentiated permit exchange rates would reduce the likelihood of high ozone levels in areas that would not otherwise be in attainment with ambient air quality standards and where the associated incremental damages to human health and welfare are relatively high.

The literature suggests that the first two of these conditions are feasible. The broader research for this project, which is out of the scope of this dissertation, will eventually use weather and atmospheric chemistry modeling to address these two conditions and the last in detail. This dissertation considers the third condition – that there is sufficient short-term flexibility to reduce NO_X emissions appreciably given realistic assumptions about the electricity markets and physical network in which they operate (Chapter 5).

3.4. Summary

This chapter reviewed the environmental economics literature to show that it is theoretically possible for a differentiated cap-and-trade program to efficiently address the time and locational impacts of environmental problems and that this theory is applicable to the impacts of NO_X emissions on ozone pollution. The environmental economics literature also discusses the implementation challenges associated with differentiated permit trading programs and potential second-best solutions. Sections 3.1.2 and 3.1.3 reviewed this literature.

Despite the scientific motivations for differentiation discussed in Chapter 2 and the economic justifications discussed in this chapter, regulators have only implemented un-differentiated cap-and-trade programs for NO_X emissions from stationary sources in the Eastern United States. The studies that review the effectiveness of these regulations offer little evidence in support of or against the use of differentiation (Section 3.2 reviewed these studies). The reasons for the lack of differentiation in the NO_X cap-and-trade programs are the implementation challenges. This chapter offered evidence that some of these could now be overcome – largely because of experience with cap-and-trade programs and because of technical improvements in air quality forecasting and modeling, in emissions monitoring, and wholesale electricity markets. The final section in this chapter discussed one conception of a potentially implementable, differentiated NO_X cap-and-trade program for ozone.

Chapter 4 – An Example Benefit of Differentiation from Coal Power Plants in the Eastern U.S.

The previous chapters presented the scientific and economic justifications for a differentiated NO_X cap-and-trade program to address ozone pollution. The scientific justification is that the relationships between the weather and biogenic VOC emissions mean that NO_X emissions at different times and locations affect ozone concentrations differently. The economic rational for a differentiated regulation is that it could achieve the same amount of environmental improvement while requiring fewer total NO_X reductions if the NO_X reductions it did require were those that directly mitigated ozone formation in areas and times with high ozone concentrations. This chapter presents an example of potentially ineffective, and therefore inefficient, NO_X abatement that occurred under the seasonal, un-differentiated cap-and-trade programs for stationary sources that have been operating the Eastern U.S. since 1999.

The seasonal NO_X cap-and-trade programs cap the total emissions from affected stationary sources between May and September each year (the "ozone season"). The seasonality of the programs is a crude recognition of the dependence of ozone formation on hot, sunny weather. But within the ozone season, the operators of power plants could make emissions reductions at any time. Section 4.4 discusses three incentives that could have affected the timing of NO_X abatement strategies of the operators of coal power plants within the ozone season. The NO_X emission characteristics of some coal power plants could make it less costly to reduce NO_X emissions during the early parts of the ozone season (e.g. May and June) because electricity demand and therefore power prices are typically lower during this period. Power prices are typically lower at the beginning of the ozone season because electricity demand is lower do to cooler weather. But the cooler weather also means that high ozone concentrations are less likely. If the incentives discussed in Section 4.4 caused the operators of generating units to reduce emissions more during the cooler parts of the summer it could mean that fewer reductions were made when they were needed later it the summer. It could also mean that the reductions made in the early summer were ineffective and therefore an example of costly over-compliance made because of the lack of differentiation in the seasonal cap-and-trade programs.

Section 4.5 uses historical data on the NO_X emissions and generation of coal power plants and on electricity and fuel prices to simulate counterfactual NO_X emissions (what emissions might have been without the cap-and-trade program) for two generating units. This analysis suggests that at least some generating units reduced emissions more during the earlier parts of the ozone season – a behavior consistent with the incentives discussed in Section 4.4. Aggregate analysis of the NO_X emissions from coal power plants in the Northeastern U.S. in 2002 and 2005 suggests that the incentives may have had a slight overall effect on the abatement decisions for more generating units.

The analysis in this chapter provides an example of a potential benefit of a differentiated cap-and-trade program: it could reduce the quantity of ineffective NO_X reductions compared to the un-differentiated programs and it could ensure that reductions occurred when they were needed. The NO_X abatement behavior discussed in this chapter was individually rational and legal but, based on the observation that ozone formation is more likely during the later parts of the summer, it may not have reduced the most harmful ozone concentrations.

In order to motivate the analysis in Sections 4.4 and 4.5, Section 4.1 provides background information on the OTC and SIP Call NO_X Budget Trading programs in addition to that provided in Chapters 2 and 3. Section 4.2 discusses the NO_X emission characteristics of coal-fired boilers and Section 4.3 discusses the characteristics of NO_X control technologies that coal power plants have used in the regions covered by the seasonal cap-and-trade programs. These details are necessary because they motivate the incentives discussed in Section 4.4.

Sections 4.2 and 4.3 also suggest that the complex NO_X emission characteristics of coal power plants make a market-based regulation preferable to a prescriptive regulation when addressing NO_X emissions from power plants. The heterogeneity in the NO_X emissions characteristics of power plants is too great to be addressed efficiently by a regulation that uniformly mandates technologies or emission rates. Empirical data show that the relationships between NO_X emissions and boiler technologies, NO_X control technologies, efficiency, and level of utilization vary greatly between different coal-fired generators. In addition, NO_X emissions often increase nonlinearly with output and NO_X emission *rates* increase as a function of output (Sections 4.2 and 4.3). Even if all NO_X emissions impacted ozone formation equally, it would not be practical for regulators to determine the most cost-effective control strategies.

4.1. The OTC and SIP Call NO_X Budget Trading Programs

4.1.1. The OTC NO_X Budget Program (1999-2002)

The Ozone Transport Commission (OTC) NO_X Budget Program cap-and-trade program ran from 1999 through 2002.[87] The EPA assisted the OTC states in the implementation and design of the OTC NO_X Budget Program and it was largely based on the Acid Rain Program's cap-and-trade program for SO_2 (Burtraw *et al.* 2005). The cap-and-trade program was seasonal in that it only capped NO_X emissions in the ozone season (May 1st through September 30th) – reflecting the fact that hot, sunny weather drives

---

[87] The states covered in the OTC NO_X Budget Program were CT, DC, DE, MA, MD, ME, NH, NJ, NY, PA, RI, VT.

ozone formation. The program's capped total ozone season NO_X emissions from the participating sources at 219,000 tons in 1999 (and 143,000 tons in 2003) and the baseline emissions were 490,000 tons in 1990. The sources were power plants and large industrial boilers[88] and the EPA gave each an initial allocation of permits.[89] At the end of the season each source was required to hold enough allowances to cover its emissions.

Within the ozone season, for the purposes of compliance, it did not matter when the abatement occurred or which sources provide it, only that the total emissions within the period fell under the cap and that each source held sufficient allowances. Allowances that the sources did not use in one ozone season could be banked for use in following seasons. If the total banked allowanced exceeded 10 percent of the total cap, the program limited the number of banked allowances that could be withdrawn from each source's account on a 1-to-1 basis, requiring the rest to be withdrawn at a 2-to-1 ratio. This provision was called "flow control" (OTC 2003).

Some measures indicate that the OTC NO_X Budget was successful. In particular: the cap was met each year, the cap was stringent enough to cause reductions of 60% from 1990 levels in ozone season NO_X emissions (OTC 2003 pg. 6), trade did not result in significant shifting of emissions geographically and the shifting that did occur was in the direction that would improve air quality (Swift 2004, OTC 2003 pgs. 9-10), and sources found it preferable to participate in the program by reducing emissions or buying permits rather than by seeking exemption or delay via litigation. The allowance markets functioned properly as economically significant trades occurred between unaffiliated sources (Burtraw 2005 pg. 43, OTC 2003 pg. 14).

---

[88] The program applied to fossil fuel fired boilers with a maximum rated heat input capacity of 250 mmBTU/hour or greater and to electric generating units with a rated output of 15 MW or more. Some states chose to include smaller units as well. See the EPA's, "Ozone Transport Commission (OTC) NO_X Budget Program Overview" at http://www.epa.gov/airmarkets/otc/overview.html.

[89] The allocation process differs by states and is fairly complicated. See Burtraw *et al.* 2005 for a summary.

Farrell 2003 has noted, however, that summertime NO_X cap-and-trade programs may not be well suited to tackle ozone episodes because reductions can occur at any time from any sources within the summer and region. However, reviews of the OTC and SIP Call NO_X Budget Trading Programs found "very little" state-level shifting of emissions to concentrated areas (Swift 2004 and Farrell 2003). Farrell (2003) found that the OTC NO_X Budget lowered average and peak emissions in equal proportions between 1998 and 2000. An EPA and OTC analysis showed that under the OTC NO_X Budget both daily total emissions and daily peak emissions declined since 1997 (OTC 2003 pg. 8). Swift (2004) found, in addition, that the largest NO_X emission sources abated the most under the OTC program.[90]

The major indicator suggesting that the OTC NO_X Budget Program could not solve the ozone problem in the Eastern U.S. was simply that many areas remained unable to attain the ozone NAAQS after the program ran for four years and led to substantial NO_X reductions (EPA 2003 pg. 18).

4.1.2. The SIP Call NO_X Budget Program (2003-Present)

Post 2002, the major action taken to remedy this continued nonattainment of ozone NAAQS in the Eastern U.S. was the implementation of a geographically extended OTC NO_X Budget program called the State Implementation Plan (SIP) Call NO_X Budget Trading Program.[91] After the Clean Air Act Amendments of 1990, many Northeastern states in the OTC region petitioned the EPA administrator under Section 126 of Title I to require NO_X reductions from states in the Midwest and Southeastern United States. The Northeastern States argued that emissions from these other states contributed significantly to their inability to attain the ozone NAAQS.

---

[90] Ellerman offers an explanation of similar findings for the Acid Rain Program: the largest sources are the least expensive sources of abatement, on a per-ton basis, when there are large capital costs associated with installing emission control technologies (Ellerman 2004, pg. 86).

[91] The program now involves the states AL, CT, DC, DE, IL, IN, KY, MA, MD, MI, NC, NJ, NY, OH, PA, RI, SC, TN, VA, WV (Phase 1); GA, MO (Phase 2).

In 1998, the EPA promulgated the NO_X SIP Call. The rule called on states to revise their SIPs, which the CAA requires them to submit as plans for how they will attain the NAAQS. Under the SIP Call, the EPA assigned each state an annual budget for NO_X emissions from large stationary sources during the ozone season. The states could comply with the SIP Call by reducing emissions as they chose or by participating in the cap-and-trade program that effectively extended the OTC NO_X Budget Trading Program. All states chose to comply with the SIP call by participating in the SIP NO_X Budget Program. The program is very similar to the OTC NO_X Budget Trading Program, but extends to additional states.

4.2. NO_X Emissions from Coal-fired boilers

In power plants and other boilers, the primary formation mechanism for nitrogen oxides is the high temperature fixation reaction of nitrogen and oxygen that occurs in high-temperature zones of the furnace (creating "thermal NO_X"). High temperatures dissociate atmospheric nitrogen (N_2) and oxygen (O_2) in combustion air, or in excess air in the combustion zone, into atomic nitrogen and oxygen, N and O. The stoichiometric ratio in the combustion zone (the air-to-fuel ratio; more air means more NO_X), the flame temperature, and the firing rate (the length of time the air is exposed to peak temperatures) determine the rates of the two reactions that form nitric oxide (NO), the principle component of nitrogen oxide ($NO_X = NO + NO_2$) emissions from combustion.[92]

Fuel nitrogen also contributes to NO_X emissions from coal-fired boilers. But its contribution is less important at high combustion temperatures because the fixation of atmospheric nitrogen dominates. Another mechanism forms "prompt" NO_X through an intermediate reaction with hydrogen cyanide (HCN). The formation of prompt NO_X does not depend much on temperature and forms under fuel-rich conditions, conditions that

---

[92] The two reactions are called the Zeldovich equations: $N_2 + O \rightarrow NO + N$ and $N + O_2 \rightarrow NO + O$.

reduce the formation of thermal NO_X.[93] Figure 4-1 shows the temperature dependence of thermal, fuel, and prompt NO_X formation for a coal-fired boiler (from U.S. EPA 1991).

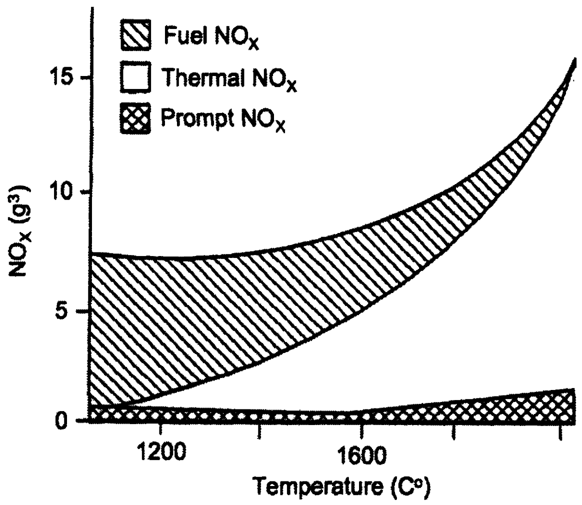

Figure 4-1 Temperature dependence of NO_X emissions from coal-fired boilers (from U.S. EPA 1991).

The different mechanisms of NO_X formation combined with difference design-specifications of coal boilers cause the NO_X emission rates, even among uncontrolled coal power plants to vary considerably. Thermal NO_X emissions increase with the availability of oxygen, the temperature of the combustion gas, and the length of time oxygen and nitrogen are exposed to peak flame temperatures. The latter two increase at higher levels of boiler utilization. NO_X emission rates tend to be higher in larger furnaces and those with high heat release rates. High heat release rates are associated with higher peak combustion temperatures that increase NO_X emissions.[94] Generating units with quicker, more intense burning have higher NO_X emission rates; these characteristics

---

[93] U.S. EPA 1991

[94] Larger furnaces have lower furnace surface-to-volume ratios. The results of a lower surface-to-volume ratio is higher furnace temperatures and other changes to the combustion process that tend to cause higher NO_X rates.

correspond to a more complete mixing of fuel and combustion gasses.[95] High levels of excess air also cause high NO_X emissions. Units with many burners close together also tend to have high NO_X emissions.

Conversely, generating units that operate on low levels of excess air, a large amount of internal recirculation of combustion gas, and slower mixing processes for fuel and air typically have lower NO_X emissions. Just as a high heat release (firing) rate is associated with high NO_X, a high thermal quenching rate (the rate of removal of the combustion-created heat) is typically associated with lower peak combustion temperatures and therefore means lower NO_X rates. Tangentially fired boilers typically have the lowest NO_X emissions among coal-fired units because the entire furnace acts as a burner so the fuel to air ratio at precise points does not matter as much as in other boilers. Also, tangentially fired units operate with low excess air as do fluidized bed boilers, which also have low NO_X emissions.

4.3. NO_X Control Technologies for Coal-fired Boilers[96]

Power plants and other industrial boilers can control NO_X emissions in a number of ways. These range from simple operational changes to combustion modifications to post-combustion ("end of pipe") controls. The specific characteristics of the boiler and the method, or multiple methods, employed determine the potential reduction in NO_X emissions from these options for a particular source.

4.3.1. Simple Operational Controls

Fuel switching is one method by which boilers burning coal or oil could reduce NO_X emissions. Coal and oil produce the most NO_X emissions and natural gas the least.

---

[95] Boilers with higher turbulence also have higher NO_X emissions. The mixing of air and fuel in the primary combustion zone creates more turbulence. Yellow hazy flames are associated with low turbulence and blue, well-defined flames with high turbulence.

[96] This discussion summarizes the detailed information in U.S. Army Corp 1998 and U.S. EPA 1994.

The reason for the differences between fuels relates to the amount of fuel-bound nitrogen and to combustion temperatures. But switching fuel is not always an option because of economic or technical considerations.

Load reduction, or decreased utilization, is another simple operational change that reduces NO_X emission rates. Load reduction not only decreases fuel use, and therefore emissions, but also decreases heat release rate and furnace temperature. Lower furnace temperatures decrease the *rate* of NO_X formation and load reduction therefore decreases the NO_X emission *rate* of a coal-fired boiler.

The reductions in NO_X emission rates from load reduction are not trivial. Technical publications suggest that a load reduction of 25 percent can cause a decrease in NO_X emission rate of around 25 percent in some pulverized coal, tangentially fired boilers (e.g. U.S. Army Corp 1988). One tangentially fired unit in a coal power plant in the Eastern U.S., for example, had an average NO_X emission rate of 0.36 lbs/mmBTU in 2002. For hourly generation levels between 860 and 870 MW, however, its average emission rate was 0.45 lbs/mmBTU and for load levels 25% lower (from 646 to 652 MW), its average emission rate was 0.29 lbs/mmBTU. This is a reduction in emissions *rate* of 35% from a 25% reduction in output. Previous studies have suggested that output reduction was one of the chosen compliance strategies under the OTC NO_X Budget Program, although they do not discuss the magnitude of NO_X reductions achieved with this abatement strategy (e.g. Burtraw and Evans 2004).

Figure 4-2 shows scatter plots of hourly NO_X emission rate versus output for this unit (left graph, Plant 6094) and another tangentially fired coal unit (in Plant 1571) in the ozone season of 2002. The graphs demonstrate that the plants' emission rates increase with output and that the relationships between NO_X emissions and output are different for the two plants, despite the fact that they employ the same technologies (tangentially fired boilers with low NO_X burners). Differences in size, age, and design and operational characteristics likely determine the difference in the NO_X characteristics of these units. One noteworthy point, however, is that the technical discussion suggests that NO_X emission rate increases with the size of boilers, but the unit in Plant 6094 has twice the capacity of that in Plant 1571 and its NO_X rate is generally lower for a given level of output.

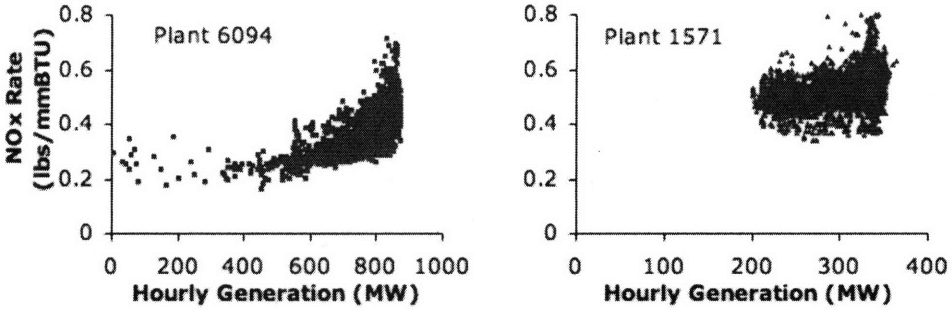

Figure 4-2 Scatter plots of hourly emission rate versus output for two tangentially fired coal power plants in the Eastern U.S. between May 1st and September 30th of 2002.

In addition to decreasing a plant's NO_X emission rate, reductions in output can also decrease a plant's efficiency. Whether or not this occurs also depends on boiler characteristics and initial load level. Figure 4-3 shows plots of hourly heat rate (inverse efficiency) versus hourly generation for the same two plants. The heat rate of Plant 6094 (left graph) increases with *decreasing* output, indicating that load reductions decrease its efficiency. The heat rate of Plant 1571 (right graph) has the opposite relationship with output; its efficiency increases slightly with decreases in output.

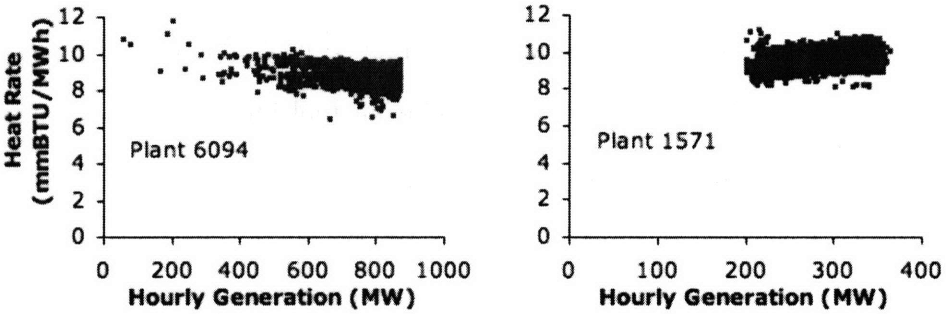

Figure 4-3 Scatterplots of hourly heat rate versus generation for two tangentially fired coal power plants in the Eastern U.S. between May 1st and September 30th of 2002.

Again, the relationships between output and efficiency are not solely results of technology as these plants employ the same technologies. The exact characteristics that cause these differences are not clear but could be related to size, age, and design and operational choices.

4.3.2. Combustion Modifications

There are a number of NO_X control methods that involve altering the combustion process. These methods either reduce peak gas temperatures, oxygen concentrations in the high temperature areas, the length of time combustion products remain in the high temperature areas, or a combination of these.

Reducing excess air is a combustion modification to control NO_X. Boilers need some excess air (beyond theoretical stoichiometric requirements) to complete the combustion of fuel. To decrease excess air, units must install combustion control systems and high quality fuel and air distribution systems to regulate and monitor the ratios of fuel and air used. Reducing excess air can improve boiler efficiency by reducing stack heat loss. Reducing normal amounts of excess air by about 50 percent can reduce NO_X by 15 to 40 percent depending on the initial levels (U.S. Army Corp 1988).

Flue gas recirculation is a combustion modification where the slightly cooled combustion gases are returned to the furnace to reduce peak flame temperatures. Fuel reburning is another method that involves introducing fuel into the boiler in stages and then completing combustion with overfire air. Fuel reburn can reduce boiler efficiency by between 0.5 and 1.5 percentage points (U.S. EPA 1994).

Two-stage combustion, also called off-stoichiometric firing, is another set of combustion-altering methods to reduce NO_X emissions.[97] These methods lower attainable peak flame temperatures by enforcing gradual mixing of fuel and air and they lower the concentration of oxygen and nitrogen in the primary combustion zone. All of these methods ensure that a large amount of the combustion occurs under conditions that suppress NO_X formation. Three types of two-stage combustion or off-stoichiometric firing are overfire air, burners-out-of-service, and low NO_X burners.

Overfire air and burners-out-of-service are a type of off-stoichiomteric combustion where alternate burners are fired under fuel-rich and air-rich conditions. For pulverized coal plants this typically means using upper burners in an air-rich or air-only manner, for that reason it is often referred to as overfire air. Overfire air can also refer to the method of operating burners fuel-rich with air injected above the burners. Low NO_X burners refer to a pattern of mixing air and fuel that both dissipates heat quickly and keeps flame temperatures low.

For coal-fired boilers, two-stage combustion reduces NO_X about 35 percent on average (U.S. Army Corp 1988). Many boilers employ multiple methods of two-stage combustion, like a combination of low NO_X burners and overfire air. There is little information available on how much one might expect NO_X control methods to impact efficiency. In 1994 the EPA reported that low NO_X burners can reduce the efficiency of

---

[97] See EPA (U.S. Army Corp 1988) and EPA's, "Air Pollutants and Control Techniques: Nitrogen Oxides," at http://www.epa.gov/eogapti1/module6/nitrogen/control/control.htm.

coal-fired units by between "0.5 and 1.5 percentage points" and overfire air can reduce efficiency be between "0.4 and 0.7 percentage points" (U.S. EPA 1994).

A simple calculation demonstrates the potential impact of the tradeoff between efficiency and NO_X emissions. If a generating unit had an initial efficiency of about 34%, its heat rate would be 10 mmBTU/MWh: it could generate 100 MWh (or 341.3 mmBTU) for heat input of 1000 mmBTU.[98] In this case, a decrease in efficiency of "1 percentage point" implies a change from 34% to 33%, which also implies a change to about 331.3 mmBTU (or 97.1 MWh) of generation from 1000 mmBTU of heat input. Thus, a reduction in efficiency of 1 percentage point corresponds to an increase in heat rate of about 0.3 mmBTU/MWh for a generating unit with an initial heat rate of 10 mmBTU/MWh. Generating with the higher heat rate would increase costs by about $0.50/MWh with coal prices at $1.5/mmBTU.[99] If NO_X-permit prices were $1000/ton, a decrease in NO_X rate of 30% from 0.6 lbs/mmBTU to 0.42 lbs/mmBTU would decrease NO_X costs by $0.84/MWh.[100]

When a NO_X control technology even slightly reduces a generating unit's efficiency, it increases fuel costs to produce a given level of output. The reduction in efficiency can also effectively reduce a generating unit's capacity. This is because generating units have an "input" capacity, which is a limit on the amount of heat input the unit can handle. At this maximum level of heat input, the unit produces maximum amount of power that is the heat input divided by the unit's heat rate. If the heat rate increases, the maximum possible output (maximum gross output) of the unit decreases.

---

[98] The conversion factor is 3.413 mmBTU/MWh.

[99] The change in heat rate from 10 to 10.3 mmBTU/MWh adds (0.3 mmBTU/MWh)*($1.5/mmBTU) to fuel costs.

[100] If the NO_X rate was 0.6 lbs/mmBTU at a heat rate of 10 mmBTU/MWh, this is 6 lbs/MWh. NO_X prices of $1000/ton are equivalent to $0.50/lb. The NO_X cost without the control would be $3/MWh. With the control, and increased heat rate, the output-based NO_X rate is 4.33 lbs/MWh. This gives NO_X costs of $2.16/MWh. The difference is $0.84/MWh.

4.3.3. Post-combustion Controls

Selective Catalytic Reduction (SCR) and Selective Noncatalytic Reduction (SNCR) are two types of post-combustion controls for NO_X emissions. These systems inject ammonia or urea into the stack gas. Because of the preference of ammonia to react with NO rather than other gasses in the stack gas, its presence reduces nitrogen oxides to molecular nitrogen and water. SNCR can reduce NO_X by 20 to 60% and SCR by 75 to 90% (U.S. EPA 1994). These technologies also cause some parasitic power losses, which could effectively reduce the generating unit's capacity. In this case, however, the unit's gross generation would remain unaffected and while the net generation that it could supply to the grid would decrease.

4.3.4. Control Technologies Used in the NO_X SIP Call Region

The EPA tracks the control technologies used by sources that fall under the NO_X SIP Call in the northeastern United States. There are approximately 709 coal units (2005 data) in this region that primarily supply electricity to the grid. In the ozone season of 2005, this group of 709 coal boilers utilized about 55 different combinations of NO_X control technologies. Table 4-1 shows the most common combinations of control technologies and the number of units that used them during the 2005 ozone season. The most popular were low-NO_X burners (LNB) and overfire air. Over half of the coal units in the SIP Call region used low-NO_X burners, alone or in conjunction with other technologies. Policies in Title IV of the 1990 Clean Air Act Amendments were partially responsible for the popularity of low-NO_X burners. The policies required some coal-fired boilers to meet NO_X emission rate standards that they could generally achieve with low-NO_X burner technologies (Burtraw *et al.* 2005).

Table 4-1 Number of coal-fired units in the NO_X SIP Call region that employed various control technologies and combinations of control technologies in 2005.

Technology	Count
Uncontrolled	123
Low NOx Burner Technology (Dry Bottom only)	109
Low NOx Burner Technology w/ Closed-coupled/Separated OFA	54
Low NOx Burner Technology w/ Overfire Air	48
Overfire Air	45
Other	31
Low NOx Burner Technology w/ Separated OFA	30
Low NOx Burner Technology w/ Closed-coupled OFA	29
Overfire Air Selective Catalytic Reduction	24
Selective Non-catalytic Reduction	22
Low NOx Burner Technology (Dry Bottom only) Selective Catalytic Reduction	21
Low NOx Cell Burner Selective Catalytic Reduction	16
Selective Catalytic Reduction	16
Low NOx Burner Technology w/ Overfire Air Selective Catalytic Reduction	13
Combustion Modification/Fuel Reburning	12
Low NOx Cell Burner	11
Low NOx Burner Technology w/ Separated OFA Selective Catalytic Reduction	10

There were also twelve different boiler technologies employed by the 709 generating units. Tangentially fired and dry-bottom wall-fired boilers were most common, each representing about 35% (249 and 246 respectively) of the boilers that operated in the ozone season of 2005.

The NO_X emission rates of generating units with the same boiler technologies and NO_X control technologies can vary considerably. Figure 4-4 shows box plots that summarize the 2005 ozone season average NO_X emission rates for coal power plants in the SIP Call region. The box plots categorize 346 of the 709 units into groups according to boiler technology and NO_X control technology. The rightmost category summarizes data for all 709 units. In some cases, as for dry-bottomed wall-fired boilers with low-NO_X burners (DBWF-LNB), the variation in average NO_X rate between units with the same technologies is almost as great as that between all the units. This does not even take into account variation in the relationships between output and NO_X rate.

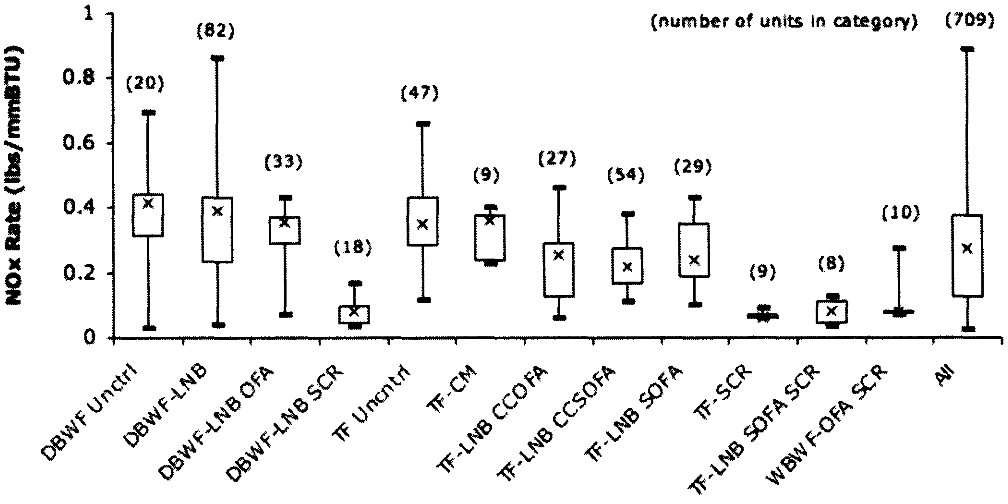

Figure 4-4 Box plots of average ozone season NO_X emission rate for coal generating units grouped into categories by boiler and NO_X control technologies. The technology abbreviations are as follows: dry-bottomed wall-fired (DBWF), tangentially fired (TF), wet-bottomed wall-fired (WBWF), low-NO_X burners (LNB), overfire air (OFA), selective catalytic reduction (SCR), combustion modification (CM), closed-coupled overfire air (CCOFA), closed-coupled separated overfire air (CCSOA), separated overfire air (SOFA), and uncontrolled (Uncntrl).

Generating units that employed SCRs tended to have lower NO_X emission rates than other coal fired units and the variation in the emission rates of these units was also smaller (labeled DBWF-LMB SCR, TF-SCR, and TF-LNB SOFA SCR in Figure 4-4). As might be expected, the median NO_X rates of units that employed control technologies were lower than those of uncontrolled units (labeled "Unctrl" in Figure 4-4); but the range in emission rates was almost as large for some categories of units with NO_X controls as for those without. For example, tangentially-fired units with closed-coupled overfire air (CCOFA) had emission rates that ranged from 0.06 to 0.46 lbs/mmBTU and uncontrolled tangentially fired units had emission rates ranging from 0.12 to 0.67 lbs/mmBTU – the minimum rates were 13 and 18% of the maximum rates respectively. Overall, the "cleanest" coal power plant had a NO_X emission rate that was about 3% of the NO_X emission rate of the "dirtiest" plant.

Some of the variation in emission rates could be a result of policies that required some coal-fired boilers to reduce NO_X emissions prior to the seasonal cap-and-trade programs. Title IV of the 1990 Clean Air Act Amendments specified two phases of NO_X emission rate standards for coal-fired electric generating units. Phase I required tangentially-fired coal boilers to obtain an emissions rate below 0.45 lbs/mmBTU and dry bottom wall-fired to obtain an emission rate below 0.5 lbs/mmBTU. Phase II required boilers to reduce their NO_X emission rates to between 0.4 and 0.86 lbs/mmBTU, depending on type. Although some of the coal-fired boilers in this sample from the Eastern U.S. appear to have average emission rates outside of these specifications, the median and even 75th percentile emission rates are about at the required levels. However, there is still considerable variation in the units' emission rates below these specifications.

The engineering documentation about NO_X emissions from coal boilers, discussed above, indicated that boiler size could be an indicator of NO_X emission rate, with larger boilers tending to have higher emission rates. For many of the technology categories, this appears not to be the case. For example, Figure 4-5 shows box plots of 2005 average ozone season emission rates for coal generating units in two technology categories grouped by heat input capacity (dry-bottom wall-fired boilers with low-NO_X burners and uncontrolled tangentially fired boilers). In both cases, the median NO_X rates of the generating units with the largest heat input capacities were lowest and the largest range of emission rates was found in the smallest size category.

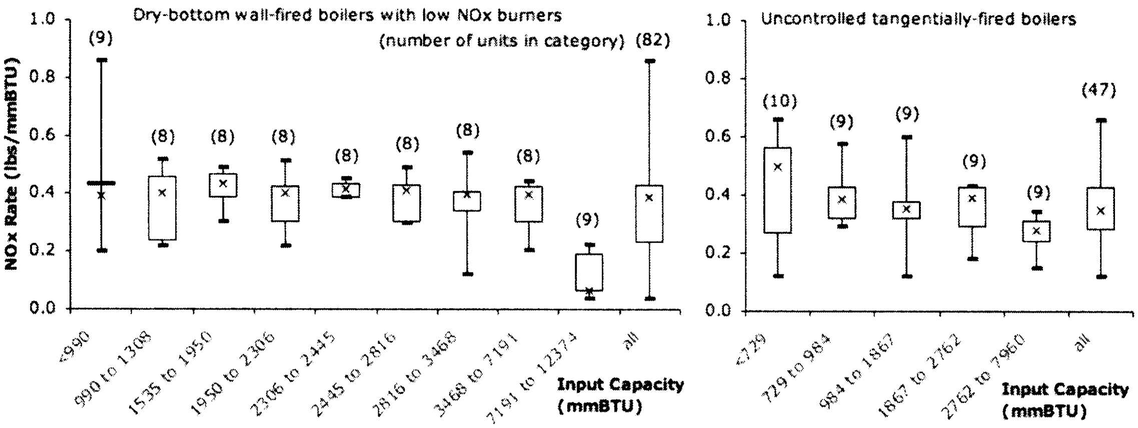

Figure 4-5 Box plots of average 2005 ozone season NO_X emission rates for generating units with dry-bottom wall-fired boilers and low-NO_X burners and for uncontrolled tangentially fired boilers in the NO_X SIP Call region. Each technology category is divided into bins based on heat input capacity.

The considerable heterogeneity in emission rates that still existed in 2005 suggests that generating units found it worthwhile to undertake different levels of abatement under the cap-and-trade programs. That is, the heterogeneity in abatement-cost between these units was (and is) not trivial and the flexibility afforded by the market-based incentive likely achieved the regulation's aggregate NO_X reductions at lower cost than would have been possible by mandating that all units achieve a particular emission rate. In addition, with so much variety in the emission characteristics of generating units with the same technologies and similar sizes it would have been impractical for the EPA to determine the most efficient control technologies or quantities of NO_X abatement for these plants.

4.4. Incentives to Reduce NO_X in the Early Ozone Season

The heterogeneity in NO_X-emission characteristics of coal power plants in the Eastern U.S. is a reason to utilize market-based incentives. But the complicated chemistry of ozone formation and, in particular, its dependence on meteorological conditions creates a challenge for the un-differentiated cap-and-trade programs now in place. The most cost-effective regulation to reduce the likelihood of peak ozone concentrations would provide stronger incentives for sources to abate emissions during the times and in the locations when ozone formation was most likely. A seasonal, but un-differentiated, cap-and-trade program does not provide extra incentives for emission reductions when ozone formation is most likely. In addition, this section discusses three incentives that could cause power plant operators to reduce NO_X emissions more when ozone formation was *less* likely under a un-differentiated cap-and-trade program.

All three of these incentives stem from the coincidence that the same weather conditions drive ozone formation and peak electricity demand.[101] Wholesale electricity markets are designed so that power plant operators have incentives (i.e. high prices) to be available and to produce more power when electricity demand is high. First, as discussed above, output reduction is a potential means to reduce NO_X emissions and it is made more attractive by the fact that the NO_X emission *rates* of some coal-fired boilers decrease with decreased levels of output. Reducing output carries a higher opportunity cost when electricity prices are high than when they are low. For this reason plant operators would prefer to utilize output reduction as a NO_X abatement strategy during cooler weather, which is also when ozone formation is less likely. Second, the use of a NO_X control technology could cause a unit's marginal costs to increase more quickly as a function of output than if it did not employ the technology. In this case, it might only be worthwhile to utilize the control technology when operating at low output levels and not at the higher levels justified by power prices observed during the hottest parts of the summer. Third, if a NO_X control technology reduced a generating unit's efficiency it can reduce its maximum output. The additional capacity is also more valuable when power prices are high.

---

[101] Hot, sunny weather provides the conditions most conducive to ozone formation. The sunlight drives the photochemical ozone formation reactions. Also, trees and other biogenic sources emit more VOCs in hot, sunny weather and high VOC levels make it more likely for NO_X emissions to contribute to ozone episodes (e.g. Fiore 2005 and Purves *et al.* 2004).

The following three subsections discuss these incentives in more detail. Following the theoretical discussion, we evaluate empirical evidence that some plant operators responded to these incentives.

4.4.1. Incentive 1: reduced generation at a given power price

The discussion of incentives for power plants to reduce NO_X emissions more during the early months of the ozone season, when power prices are low, begins from the basic concepts of profit maximization and marginal costs for power plants. Assuming a competitive market, the operator of a generating unit (i) will choose a level of output (q) in an hour (t) that maximizes profit:

$$\Pi_{it} = p_t q_{it} - \int_0^{q_{it}} mc_i(q_{it}) dq_{it}$$

where p_t is the price of power at time t and $mc_i(q)$ is unit i's marginal cost of producing q.[102] That is, the operator will choose a level of output, q, such that its marginal cost of producing q are less than or equal to the price of power:

$$mc(q_{it}) \leq p_t.$$

Baseload units with low variable costs, like coal-fired power plants, generate in most hours and when power prices are greater than a plant's marginal cost of producing its maximum output (q_{max}), it earns inframarginal rents that cover fixed costs. This margin increases with power prices and can be quite large in some hours (Appendix A gives examples). The price cap for wholesale power markets in the Eastern U.S is $1000/MWh. Prices of this level do not reflect the marginal costs of the generating unit that would provide an additional increment of power. Instead they indicate scarcity conditions when

---

[102] The price of power also depends on the generator's location in the power network because the wholesale power markets employ "locational marginal prices" (LMP). The price of power could therefore be indicated p_{it} but this aspect is ignored for simplicity. The LMP for a particular generating unit's node on the power network varies from other nodes, when congestion affects it.

the system operator cannot maintain operating reserves.[103] In this case, higher NO_X prices could not cause increases in the electricity price and would only reduce generators' profits.

Coal generating units are often large boilers that cannot be easily, or inexpensively, turned on and off. This means that even when the price of power falls below the marginal costs of doing so, most coal power plants generate at a minimum level of output (q_{min}). A boiler also has a heat input limit, which restricts its maximum output at some level (q_{max}). A generating unit's maximum output level is often referred to as its "rated capacity". A coal unit's *minimum* output level is typically between 20% and 40% of its rate capacity.

The cost of fuel dominates the variable costs of generating units. Other variable costs include operating and maintenance (O&M) costs and the costs of emissions under regulations like SO_2 and the seasonal NO_X cap-and-trade programs. Non-fuel O&M costs are small for power plants and since these data are difficult to find for individual generators or power plants; this analysis only considers fuel costs. The relationship between fuel use and output, which is called an input-output curve, is the starting point for developing marginal cost curves for generating units.

A generator's "input-output" curve, $\mathbf{G}(q)$, describes how hourly input of fuel (in mmBTU per hr) to the generator varies with its hourly output, q (in MW). The observed input-output relationships of coal power plants over their typical ranges of utilization are linear. For this reason, and for simplicity, we employ linear input-output curves.[104] Multiplying the input-output curve by the cost of fuel, p_f, in units of dollars per energy

---

[103] For example see ISO New England's Operating Procedure Number 4 (OP 4) at http://www.iso-ne.com/rules_proceds/operating/isone/index.html .

[104] The industry standard is to model the input-output relationships of a generator, i, as cubic: $\mathbf{R}_i(q) = a_i + b_i q + c_i q^2 + d_i q^3$ (e.g. Obessis and Lamont 1993). In reality though, these are often simplified to linear.

input (\$/mmBTU) then gives the unit's fuel cost as a function of power output, with units of dollars per hour: $\mathbf{C}_{\mathbf{f}\_i}(q) = p_f\mathbf{G}_i(q)$.

Assume first that the unit's operator cannot change its emission rate within the ozone season with a control technology but can only reduce output in order to reduce NO_X. Based on the previous section's discussion, this is a simplification because a NO_X control technology can impact the heat rate of a generating unit and therefore plant operators may not find it worthwhile to utilize the control technologies for the entire summer. The simplified case is relevant, however, because even with the use of a control technology generating unit operators can reduce output to obtain additional reductions in NO_X. This section examines the simplified case and the following section examines the case where the operator can alter the use of a NO_X control technology during the ozone season.

If $\mathbf{N}_i(q)$ is the relationship between the unit's hourly output and NO_X emissions (in lbs/hr) and the cost of emissions is the opportunity cost of selling permits at their market value, the cost curve for fuel and NO_X emissions takes the form: $\mathbf{C}_i(q) = p_f\mathbf{G}_i(q) + p_n\mathbf{N}_i(q)(\mathbf{I})$, where p_n is the price of NO_X (in \$/lb), and **I** is a dummy variable indicative of the ozone season. The derivative of the cost curve yields the marginal cost curve for a generator *i*, $mc_i(q)$, which explains how fuel and NO_X costs change with changes in output. (The subscript, *i,* is dropped for notational simplicity in the remainder of the discussion.) In a competitive market, a generator will produce *q* when the marginal costs of doing so are less than or equal to the price of power (*p*):

$$d\mathbf{C}(q)/dq = mc(q) = p_f\,(d\mathbf{G}(q)/dq) + p_n\,(d\mathbf{N}(q)/dq)(\mathbf{I}) \leq p.$$

The operator will run the unit at q_{max} when the price of power is above the marginal cost of doing so.[105] The unit's marginal costs are higher (by p_n $(d\mathbf{N}(q)/dq)$ during the ozone season because of the price on $\mathrm{NO_X}$ emissions; thus, inside the ozone season it would not be economic to generate at full capacity until power prices were slightly higher to reflect this additional cost.

Now consider a case that is common for coal generators: an exponentially increasing $\mathrm{NO_X}$-output relationship and a linear input-output relationship. Still assuming that the operator cannot use a control technology to change the unit's emission rate within the ozone season, this gives:

$$\mathbf{N}(q) = \mathrm{k}e^{\mathrm{m}q} + \mathrm{u} \text{ and k,m} > 0 \text{ and} \tag{1}$$

$$\mathbf{G}(q) = \mathrm{a}+\mathrm{b}q. \tag{2}$$

The generator's marginal costs are constant outside the ozone season, but increase with output during the ozone season and it will produce q when:

$$\mathrm{mc}(q) = \mathrm{p}_f\mathrm{b} + \mathrm{p}_n\mathrm{mk}e^{\mathrm{m}q}(\mathbf{I}) \leq \mathrm{p}. \tag{3}$$

Solving for the unit's output, q, as a function of the price of power, p_p, both within an outside of the ozone season yields:

$$q = \begin{cases} q_{min} & \text{if } \mathrm{p}_p < \mathrm{p}_f\mathrm{b} \text{ and } \mathbf{I} = 0 \text{ or if } \mathrm{p} < \mathrm{mc}(q_{min}) \text{ and } \mathbf{I} = 1. \\ q_{ozone}(\mathrm{p}) & \text{if } \mathrm{mc}(q_{min}) \leq \mathrm{p} \leq \mathrm{mc}(q_{max}) \text{ and } \mathbf{I} = 1 \\ q_{max} & \text{otherwise.} \end{cases} \tag{4}$$

---

[105] Assuming that a unit will produce at its full capacity when the power price exceeds its marginal cost of doing so while using these fuel-cost based marginal cost curves is somewhat of a simplification. There are reasons why a unit's marginal costs of producing at full capacity might be higher than modeled here. These reasons include the opportunity costs of participating in ancillary services markets, which essentially pay generating units to operate at less than full capacity, and transmission constraints.

If $\mathbf{N}(q)$ is exponential, then for $\mathrm{mc}(q_{min}) \leq \mathrm{p} \leq \mathrm{mc}(q_{max})$ Equation (4) becomes:

$$q_{ozone}(\mathrm{p}_p) = (\mathrm{m})^{-1} [\ln(\mathrm{p} - \mathrm{p}_f\mathrm{b}) - \ln(\mathrm{p}_n\mathrm{mk})]. \qquad (5)$$

Thus, depending on the values of the coefficients, a range of power prices may exist for which it is profitable for a unit to generate at full capacity outside the ozone season, but not during the ozone season. This creates an incentive for output reduction during the ozone season for this range of power prices, which Figure 4-6 illustrates as $\underline{p} \leq p < p_{max}$ using hypothetical marginal cost curves.

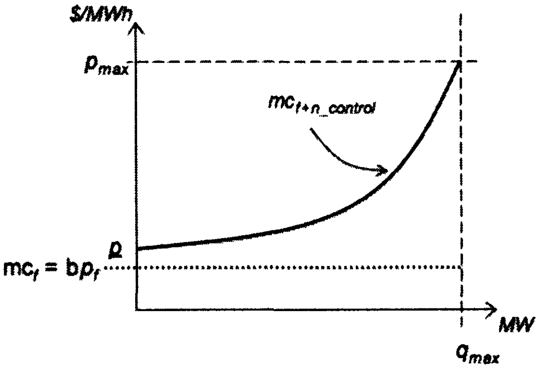

Figure 4-6 Marginal cost curves for a generating unit without a NO_X price, c_f, (i.e. non-ozone season) and with a NO_X price.

This incentive to reduce output at a given power price is similar when a generator's NO_X-output curve is quadratic. In either case, if a generator's NO_X-output curve increases nonlinearly, its NO_X emission *rate* increases at higher output levels. A plant operator has a stronger incentive to reduce output as an abatement strategy for a unit with a NO_X rate that increases with output, compared to one with an emission rate that is a constant or decreasing function of output.

4.4.2. Incentive 2: variable use of NO_X control technologies

The discussion this far assumed that a generating unit's NO_X-output relationship remained constant throughout the ozone season; it assumed that the only way a generator could decrease NO_X emissions was to decrease utilization. But at least some NO_X control technologies provide the opportunity to alter a generator's NO_X-output relationship (or NO_X emissions rate at a given q) on a time scale shorter than the five-month ozone season. Some combustion-modifying NO_X control methods involve tuning a unit's operating conditions, which can be done somewhat flexibly. In this case, the analysis of marginal cost curves suggests that it may not be worthwhile for an operator to use NO_X control technologies during periods with high expected electricity prices if the marginal cost of abating NO_X increases steeply as q approaches q_{max}.

Assuming that the use of a control technology may diminish a unit's efficiency, and therefore require more heat input for a given level of output, the marginal cost curve for a generator with the use of a control technology is:

$$d\mathbf{C}_{control}(q)/dq = p_f(d\mathbf{R}_{control}(q)/dq) + d\mathbf{V}_{control}(q)/dq + p_n(d\mathbf{N}_{control}(q)/dq)(\mathbf{I}) \qquad (6)$$

where $\mathbf{V}_{control}(q)$ represents the variable costs of operating the control technology. The variable costs of control technologies tend to vary linearly with output, $V_{control}(q) = v_{cont}q$ (e.g. Foerter and Jozewicz 2001).

If, as before, the generator's NO_X-output relationship is exponential and the input-output relationship is linear, and coefficients with the subscript *"c"* represent the change in the coefficient caused by the NO_X control, then a generator will produce q if:

$$\boldsymbol{mc(q)} = d\mathbf{C}_{control}(q)/dq = p_f(b + b_c) + v_{cont} + p_n((m+m_c)(k+k_c)e^{(m+m_c)q})(\mathbf{I}) \le p \qquad (7)$$

with

$$\mathbf{N}(q) = (k+k_c)e^{(m+m_c)q} + (u+u_c) \text{ and} \qquad (7a)$$

$$\mathbf{G}(q) = (a + a_c) + (b + b_c)q. \qquad (7b)$$

As before, the marginal costs of NO_X emissions increase as a function of q while the marginal fuel costs are constant.

If a unit's controlled NO_X emissions increased more steeply with increasing q than its uncontrolled NO_X emissions (for example, if the coefficient $m_c > 0$), the unit's marginal costs of generating q while utilizing NO_X controls could be higher than if it did not use the controls. Then for high p, the loss of margins from decreased utilization could outweigh the savings from using a control technology.

A schematic of this possibility is sketched in Figure 4-7. The figure shows hypothetical marginal cost curves for uncontrolled ($mc_{no\_control}$) and controlled ($mc_{control}$) NO_X emissions. For $p < p_{int}$, controlling NO_X emissions with is worthwhile. But, for $p > p_{int}$ the generator could economically produce more power if it did not control NO_X emissions. For prices $p > p_{max}$ it will always be economic for this generator to produce at maximum capacity, but the operator will prefer not to employ the control technology if the shaded area B is larger than the area A.

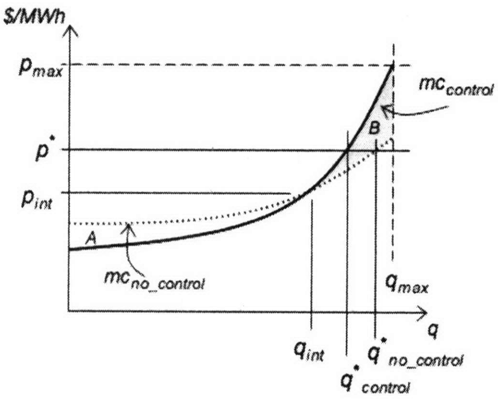

Figure 4-7 Sketch of marginal cost curves for the case when a generating unit's marginal cost with NO_X control increases more steeply than without a NO_X control technology.

At some intermediate price p^*, with $p_{int} < p^* < p_{max}$, the operator would prefer not to control emissions if the opportunity cost from producing less power plus the cost of the control technology and any heat rate penalty outweighed the savings from controlling NO_X. This occurs when:

$$p^*(q^*_{control}) - \int_0^{q^*_{control}} c_{control}(q)\,dq < p^*(q^*_{no\_control}) - \int_0^{q^*_{no\_control}} c_{no\_control}(q)\,dq.$$

The technical requirements of control technologies suggest that it is likely that a plant operator cannot turn on and off a unit's NO_X controls on an hourly basis, but could alter use of controls on a time frame of a month or a few weeks. In this case, a plant operator with a generating unit that has marginal cost curves like those shown in Figure 4-7 might choose to control NO_X only during periods when the expected power price does not exceed p^*.

4.4.3. Incentive 3: NO_X control technologies reduce capacity

Another reason that a unit's operator might prefer to use control technologies only in the early summer is that when NO_X control technologies reduce efficiency (increase heat rate) they also reduce maximum possible generation. The additional capacity, or higher maximum hourly generation, is more valuable when power prices are high. A unit's heat input capacity limits the amount of fuel that its boiler can burn in an hour. The efficiency of the unit determines the fraction of the hourly heat input that it can covert to output, or generation.[106] If the use of a control technology reduces a generating unit's efficiency, it would not be able to generate as much power from its maximum heat input.

This section discussed three reasons that power plant operators might prefer to abate NO_X emissions more when power prices were relatively low. First, output reduction as a NO_X abatement strategy is clearly less costly when power prices are low. Second, if a generating unit's marginal costs increase more steeply with a NO_X control technology than without it the operator might prefer to reduce emissions at lower power

---

[106] As mentioned above, a typical efficiency of a coal boiler is about 34%, which corresponds to a heat rate of 10 mmBTU/MWh.

prices. Third, if a NO_X control technology increases a generating unit's heat rate it can decrease its effective capacity, which is more costly when power prices are high.

4.5. Abatement under OTC and SIP Trading Programs

Historical emissions and generation data for coal power plants in the OTC and NO_X SIP Call regions provide evidence in support of the theory that some power plants have incentives to reduce emissions more when power prices are low under a un-differentiated cap-and-trade program. This section uses hourly data to assess whether the three incentives discussed in Section 3 affected the generation and NO_X emissions of three generating units. The analysis of the three example units suggests that at least some plant operators reduced the output of coal generating units as a result of the seasonal NO_X cap-and-trade program and that they reduced output and NO_X emissions more during periods with lower power prices. Weekly data on generation and emissions from power plants in the New England and PJM power systems were then used to estimate whether more NO_X abatement occurred in 2002 and 2005 during the beginning of the ozone season, when power price tend to be low, than in the later parts of the ozone season.[107] The estimates suggest that weekly abatement decreased slightly throughout both the 2002 and 2005 ozone seasons. For example, the average weekly NO_X reductions in June in 2002 were about 1930 tons while those in August were about 1620 tons.

Ozone formation is most likely during hot, sunny weather and ozone episodes are most frequent in July through September in the Eastern United States.[108] The observation of slightly more abatement in the earlier months of the ozone season (May and June) is

---

[107] Due to air conditioning demand, electricity demand and prices tend to be highest in the months of July and August (see Appendix A).

[108] EPA air quality monitoring data also suggest that exceedances of the ozone standards were more likely in the late summer. There were 25 days on which ozone concentrations exceeded the ozone standard in the New York City area during the summer of 2005 and that 22 of these occurred in the months of July, August and September. In the District of Columbia nonattainment area, all 6 days with exceedences were in August. In the Philadelphia area, 9 of 13 days with exceedances occurred in July, August and September. For these reason and for convenience, we split the ozone season into the "early" portion (May and June) and the "late" portion (July through August).Calculated from EPA Continuous Emissions Monitoring data at http://cfpub.epa.gov/gdm/.

evidence of slightly perverse behavior that could have been caused by the incentives discussed in Section 4.4. However, it is difficult to say whether additional reductions of approximately 300 tons per week in July or August might have prevented ozone episodes. Answering this type of question would require atmospheric chemistry modeling, but 300 tons was only a small portion (5%) of the total weekly NO_X emissions of about 6400 tons in July through September of 2002.

The remainder of this section discusses evidence that the incentives discussed in Section 3 affected the generation and NO_X emissions of example generating units (Sections 4.1 through 4.3 discuss the incentives in the same order as Sections 3.1 through 3.3). Section 4.4 then discusses the estimation of a counterfactual for emissions from coal power plants in the New England and PJM power systems for 2002 and 2005 and compares the estimates to the base case year of 1998 before the OTC NO_X Budget Program began.

4.5.1. Incentive 1: reduced generation at a given power price

Section 4.4.1 suggested that if a generating unit's NO_X emissions increased nonlinearly with output, a price on NO_X emission could then cause its marginal cost curve to increase nonlinearly as well. In this case, a range of power prices may exist for which the unit's operator would prefer to produce at full capacity when the NO_X price was not effective but at a lower level when it was.

Nonlinear relationships between NO_X emissions and generation (NO_X-output curves) are often observed for coal generating units. However, the relationships between heat input and generation for coal units tend to be linear. For example, Figure 4-8 shows scatter plots of hourly heat input and NO_X emissions versus output for coal units in three power plants in the ozone season of 2002.[109] The data are for two 913 MW units in Plant

---

[109] Appendix B discusses curve fitting results.

6094, two 364 MW units in Plant 1571, and one 626 MW unit in Plant 1573. The units in Plants 6094 and 1571 are dry bottomed wall-fired boilers that utilized low NO_X burners in 2002. The unit in Plant 1573 is a tangentially fired boiler that employed low NO_X burner technology and closed-coupled, separated overfire air to control it NO_X emissions. These units' NO_X-output relationships suggest that a small decrease in utilization disproportionately decreases their NO_X emissions, especially at high initial levels of output.

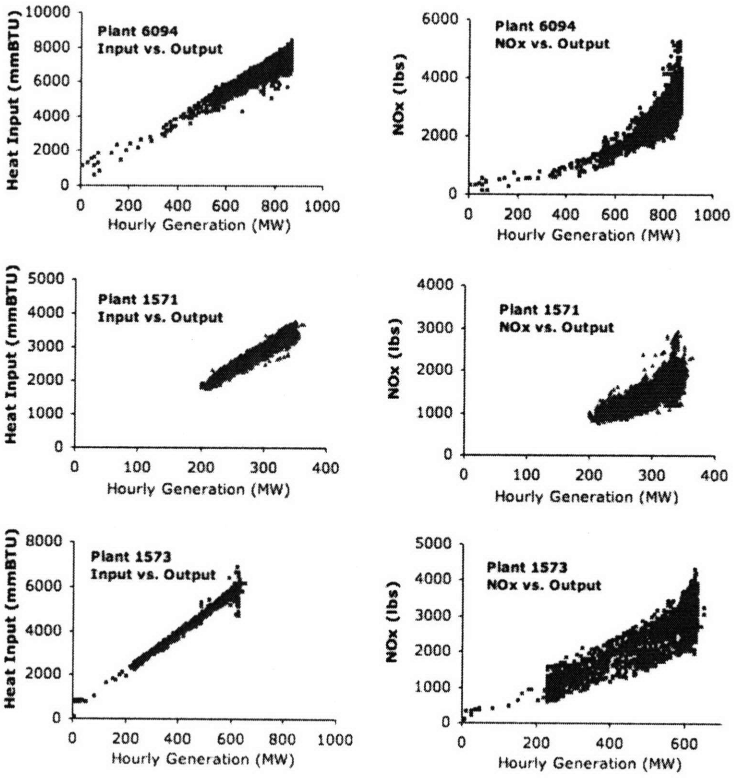

Figure 4-8 Scatter plots of hourly heat input versus generation (input-output curves) and NO_X rate versus generation for two coal-fired power plants in the ozone season of 2002.

Linear and exponential fits to the data in Figure 4-8 produced estimates of the input-output and NO_X-output curves for these units. The input-output curves were constructed by fitting linear curves to the hourly data using ordinary least squares. A dummy variable for the ozone season months was also included. The hourly data are from the EPA's Continuous Emissions Monitoring System (CEMS). For each unit, the hourly data were used to estimate the coefficients on output (q) in the relationship in Equations 7b:

$$HI = a + a_c{*}OZONE + bq + b_c(q{*}OZONE) + \varepsilon.$$

Where HI is the hourly heat input (mmBTU), q is the unit's hourly gross output (MW), OZONE is a dummy variable that equals one during the ozone season, and ε is the error term.

The NO_X-output curves were constructed similarly, but the data were transformed to obtain an exponential fit like that described in Equation 7a. For each unit, the hourly data were used to estimate the following relationship, where $\ln(NO_X)$ is the natural logarithm of the unit's hourly NO_X emissions:

$$\ln(NO_X) = k + k_cOZONE + mq + m_c(q{*}OZONE) + \varepsilon.$$

The estimates of the coefficients (discussed in Appendix B) were then combined with historical data on NO_X prices and coal prices in order to estimate marginal cost curves for the example units inside and outside of the ozone season based on Equation (7):

$$\boldsymbol{mc(q)} = p_f(b + b_c) + p_n((m+m_c)(k+k_c)e^{(m+m_c)q})(\mathbf{I}) \tag{7}$$

Figure 4-9 shows marginal cost curves estimated with the input-output and NO_X-output and weekly historical fuel and NO_X price data for three of the above units in 2002 (P6094U3, P1571U1, and P1573U2).[110] These simple curves do not include variable operating costs for the boilers or control technologies, which typically vary linearly with q and would shift the marginal cost curves upward.

---

[110] The Energy Information Association publishes monthly data on the cost of coal delivered to the power plants in each state and spot prices were obtained for Big Sandy Barge Low Sulfur Coal (Bloomberg ticker COALBGSD). The results discussed here use the spot price data, but the results were similar and only changed the threshold prices by about \$1 to 2/MWh for these units. NO_X prices were fairly stable over the summers of 2002 and 2005, about \$900/ton and between \$2000 and \$2700/ton respectively. Weekly NO_X price data from the Bloomberg OTC NO_X allowance price series (EMITNOXC) were used to create the marginal cost curves.

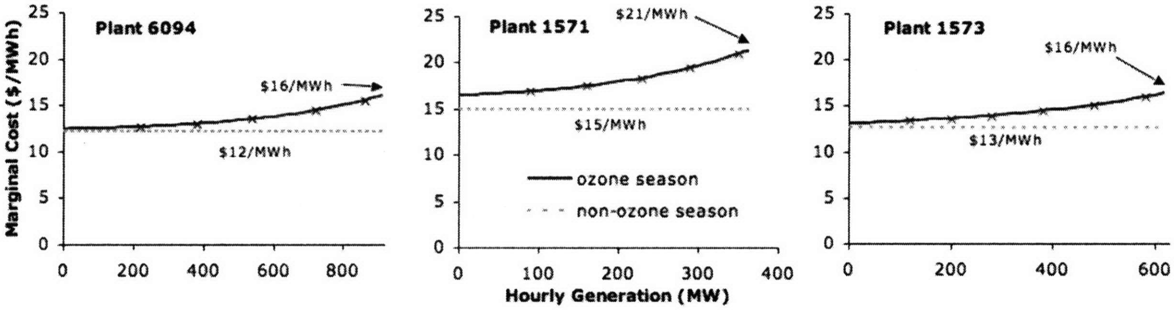

Figure 4-9 Estimated marginal cost curves for coal generating units in Plants 6094, 1571, and 1573 in 2002.

The curves yield estimates of the ranges of power prices for which the units' operators would prefer to generate slightly less power during the ozone season without the NO_X price. These ranges are labeled on the graphs in Figure 4-9 for the three units. Empirical evidence suggests that the operators of the three example units did respond to incentives for reduced output in the ozone season.For example, Plant 1573's marginal cost curves inside and outside the ozone season suggest that the operator might have decreased output inside the ozone season for prices ranging between about $13 and $16/MWh. Given power prices in this range during 2002, the mean hourly generation for P1573U2 was 280 MW in the ozone season and 400 MW outside the ozone season. The unit was also much more likely to have generated near its maximum capacity outside the ozone season when prices were in this range:

$$P_{non\text{-}ozone}(q \geq 0.9q_{max} \mid 14 \leq p < 18) = 0.08 \text{ and}$$

$$P_{ozone}(q \geq 0.9q_{max} \mid 14 \leq p < 18) = 0.00.$$

Table 4-2 shows the ozone season and non-ozone season mean generation for the three example units in each unit's relevant price range (the lowest price range shown for each unit) and in higher price ranges of equal width for comparison. The mean ozone season generation was less than that in the non-ozone season for all three units in the relevant range (and the hypothesis that the difference between the two means for each unit was zero could be rejected in all cases). The difference between the ozone and non-ozone season means decreased at higher prices ranges for all units.

Table 4-2 Ozone season and non-ozone season mean hourly generation for the three example generating units.

	Price Range	**Mean Hourly Generation**		
		Ozone	Non-Ozone	Difference*
Plant 6094	**12<p<=16**	773	785	12 *(2.6)*
	16<p<=20	804	810	5 *(1.7)*
	20<p<=24	821	819	-2 *(.40)*
Plant 1571	**15<p<=21**	259	271	12 *(7.2)*
	21<p<=27	305	318	13 *(6.2)*
	27<p<=33	317	320	3 *(1.2)*
Plant 1573	**13<p<=16**	280	400	120 *(17)*
	16<P<=19	401	342	-59 *(13)*
	19<P<=22	561	575	14 *(2.2)*
	($/MWh)		*Hourly MW*	

* T-statistic (absolute value) in parathenses; a t-statistic with absolute value greater than 1.96 indicates that we can reject the hypothesis that the ozone season and non-ozone season means are equal (5% level of significance for the 2-sided t-test).

The observation that these generating units generated less at a given price inside the ozone season suggests that the operators of these generating units may have reduced output because of the NOx price. Transmission network congestion could also affect whether a unit generates at full capacity, but network congestion is unlikely at lower levels of electricity demand, which corresponds to prices in the relevant range, and is as likely to occur both within and outside of the ozone season for similar levels of demand and price.

Another test of whether or not the NO_X price caused units to reduce output during the ozone season is a counterfactual that estimates what the generation of these units would have been without the NO_X regulation. The following section discusses observed NO_X emissions and uses counterfactuals for generation, heat input, and NO_X emissions to analyze the behavior of these plants.

4.5.2. Incentive 2: variable use of NO_X control technologies

This section evaluates empirical evidence that the operators of the two of the example generating units varied the use of the units' control technologies during the ozone season. Counterfactual estimates of generation, heat input, and NO_X emissions were used to estimate the units' NO_X abatement during the ozone season of 2002 and the portion of this abatement that was due to output reduction versus the use of NO_X control technologies. Both units reduced emissions more during the months of May and June in comparison to July through September. However, the units did not appear to greatly vary their use of NO_X control technologies throughout the summer.

One indication that a power plant operator may have used a control technology to abate NO_X emissions is the observation of a lower ozone-season emission rate. However, as discussed earlier, if a unit's NO_X emission rate decreases with level of output then an observed decrease in NO_X emission rate is not necessary indicative of the use of a control technology or even of a response to the NO_X cap-and-trade program. Figure 4-10 shows the average NO_X emission rates (lbs/mmBTU) of the three example units by week for 2002. The emission rate of P6094U3 was only slightly lower in the ozone season than outside of it. The emission rates of the other two plants were lower during May and June (weeks 18 through 25) than in July through August (weeks 26 through 40).

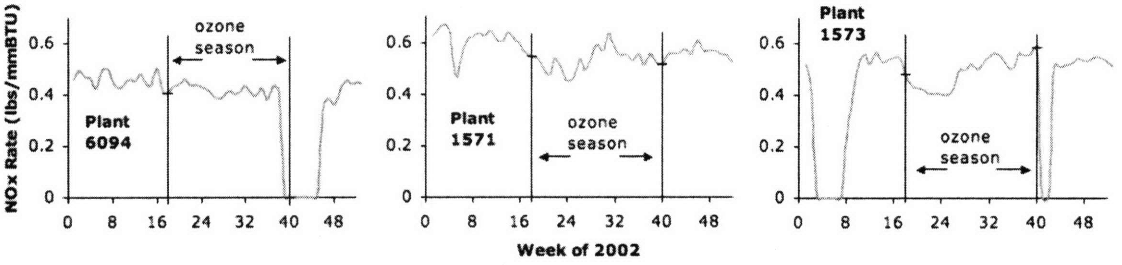

Figure 4-10 Weekly average NO_X emission rates for coal units in Plants 6094, 1571, and 1573 for 2002.

Two questions arise from the observation that at least some of the generating units under the NO_X cap-and-trade programs had lower emission rates during the "early" part of the summer and less so in the "late" part of the summer. First, what physical relationship or change caused the observed decreases in NO_X rates? Decreased output, a NO_X control technology, or a combination could have caused lower emission rates. Second, to what extent was the cap-and-trade program the motivation for the physical changes? The NO_X price and the previously discussed incentives could have caused the plants' operators to abate, but lower electricity prices or other problems like malfunctioning control technologies could also have been responsible.

Counterfactuals, or estimates of what the units' generation, heat input, and emissions would have been without the NO_X cap-and-trade program, are a useful tool to answer these questions. The seasonal nature of the cap-and-trade program can be used to construct the counterfactual because the NO_X price would not have affected the relationships between electricity price and generation, between generation and heat input, and between heat input and NO_X emissions during the non-ozone season months.

Constructing the counterfactuals requires understanding three relationships that a NO_X price can change. The first is between electricity price and generation. If one controls for fuel costs and assumes that the units in question rarely, if ever, set the electricity price, then generation can be predicted from electricity prices (p) and coal prices (p_f). It is reasonable to assume that the coal power plants in question rarely set the price of power because their variable costs are low. Additionally, it is difficult to determine which hours these units may have been the marginal unit in the power system. Regressions that used total electricity demand as a proxy for electricity price (because total demand is inelastic and is not affected by the generation of any particular plant) did not predict output as well and using demand as an instrumental variable for price did not change the results.

In general, the generation of generating units increases as a function of the electricity price. But, because the generating units have finite capacities, the relationship between electricity price and output is flat when prices are high enough for the units to consistently generate at or near their maximum. For this reason, piecewise linear relationships were fit for the generating units. The marginal cost curves discussed above predicted the break points for the piecewise linear curves: the prices at which the units could be expected to generate at their full capacities inside the ozone season. Slightly higher break points were used ($20/MWh for P1573U2 and $23/MWh P1571U1) in these estimates to account for additional variable costs not modeled in the marginal cost curves. The following relationship was fit for generation and power prices below and above the break point for each generating unit:

$$q_{it} = \beta_o + \beta_{1i}p_t + \beta_{2i}p_{ft} + \beta_{3i}EARLY + \beta_{4i}LATE + \varepsilon \qquad (8)$$

The dummy variables EARLY and LATE are one during May and June and July through September respectively.[111] Weekly coal price data and hourly generation and power price data were used in the regressions.[112] The counterfactual for generation in the

---

[111] Dummy variables for the "early" and "late" summer were chosen arbitrarily to represent the portions of the ozone season where power prices are generally low and ozone formation unlikely ("early") and when power prices are higher and ozone formation more likely ("late").

[112] The Energy Information Association publishes monthly data on the cost of coal delivered to the power plants in each state and weekly spot prices were obtained for Big Sandy Barge Low Sulfur Coal (Bloomberg ticker COALBGSD). The hourly historical locational marginal price data for the PJM pricing node that corresponded to the

absence of the seasonal cap-and-trade program was constructed by predicting generating using the observed electricity prices and fuel prices during the ozone season and the predicted coefficients from Equation (8), but setting EARLY=LATE=0.

The second relationship to consider for the counterfactual is the amount of fuel needed to produce a given level of output. A NO_X price can change this relationship, the input-output relationship, if controlling NO_X emissions reduces efficiency. In this case, the observed generation[113] was used to predict the required heat input (HI):

$$HI_{it} = \alpha_o + \alpha_{1i}q_{it} + \alpha_{2i}\text{EARLY} + \alpha_{3i}\text{LATE} + \varepsilon \qquad (9)$$

The counterfactual heat input was constructed by predicting HI_{it} from the counterfactual estimates of q_{it} (from Equation 8) with EARLY=LATE=0. This process estimated the amount of heat input that would have been required to generate the *counterfactual* generation given the non-ozone season relationship between heat input and output.

The final relationship estimated was between heat input and NO_X emissions. Although the NO_X emissions and generation of these units are correlated, heat input – or the burning of fuel – ultimately causes both NO_X emissions and generation. The relationship between NO_X and output, discussed in Section 4.5.1 and Appendix B, was exponential; but a quadratic fit between NO_X and heat input was most appropriate. The counterfactuals of NO_X emissions for each unit were constructed similarly to the heat input counterfactual, but with a squared term to reflect the nonlinearity between NO_X emissions and heat input:

$$NO_{Xit} = \mu_o + \mu_{1i}\text{HI}_{it} + \mu_{2i}(\text{HI}_{it})^2 + \mu_{3i}\text{EARLY} + \alpha_{4i}\text{LATE} + \varepsilon \qquad (10)$$

---

power plant were used for each unit; available from the PJM website, "Daily Real-Time Locational Marginal Pricing Files," at
http://www.pjm.com/markets/jsp/lmp.jsp. The hourly generation data for the units were from the EPA's CEMS data.

[113] Predicting HI on the predicted generation from Equation 8 did not change the estimates of the coefficients in Equation 9.

The non-ozone season relationship (EARLY=LATE=0) between output and input was used to predict the amount of NO_X emissions from the *counterfactual* values of heat input.

This process generated the counterfactual estimates of generation, heat input, and NO_X emissions. Table 4-3 shows the coefficients for the regressions in Equations 8, 9, and 10. These were the coefficients used construct the counterfactual estimates. Most coefficients were significant at the 5% level. The signs of the coefficients are as expected. For example, the signs on the coefficients of the power price (LMP) in the regressions of generation on LMP are positive for prices below the cut point; generation was expected to increase as a function of the price of power in this region. At prices above the cut point the magnitude of the coefficient on LMP is small reflecting the fact that the plant operators cannot increase the units' output above their capacities.

The coefficient on generation (q) in the regressions of heat input on generation is positive and about 9 or 10 (mmBTU/MWh). This was expected because this coefficient represents the units' heat rates (normally about 10 for coal fired boilers). Negative coefficients on the EARLY and LATE dummy variables indicate less generation for a given power price than during the non-ozone season. Similarly, negative coefficients on the dummy variables in the regressions of NO_X emissions on heat input indicate a decrease in NO_X emissions at a given level of heat input during the ozone season. Positive coefficients on the dummy variables in the regressions of heat input on output indicate a higher heat rate during the ozone season possibly due to the use of control technologies (or to ambient conditions).

Table 4-3 OLS coefficients used to construct counterfactual estimates of generation, heat input, and NO_X for two coal generating units.

		OLS Coefficient (standard error in parenthesis)					**Number of Observations**	**Adjusted R^2**
1571U1 2002		**Intercept**	**LMP**	**CoalPrice**	**EARLY**	**LATE**		
	q for p < 23	217.6** (16.56)	11.98** (0.339)	-5.04** (0.539)	-18.2** (1.892)	2.70 (1.774)	3190	0.303
	q for p >= 23	478.4** (14.12)	-0.040** (0.015)	-5.14** (0.506)	-4.61** (1.464)	-5.38** (1.178)	3531	0.049
		Intercept	**q**		**EARLY**	**LATE**		
	HI	-414** (7.97)	10.8** (0.026)		45** (3.46)	157** (2.96)	7933	0.958
		Intercept	**HI**	**HI^2**	**EARLY**	**LATE**		
	NOx	-280** (10.45)	0.687** (0.0036)	2.02E-06** (4.729E-07)	-243** (5.24)	-135** (4.55)	7933	0.840
1573U2 2002		**Intercept**	**LMP**	**CoalPrice**	**EARLY**	**LATE**		
	q for p < 20	-6.9 (56.3)	45** (0.86)	-9.3** (1.99)	-86.7** (4.98)	-72.1** (4.84)	2887	0.533
	q for p >= 20	734.0** (26.4)	0.18** (0.025)	-5.14** (0.956)	2.06 (2.56)	-1.50 (2.18)	3705	0.022
		Intercept	**q**		**EARLY**	**LATE**		
	HI	240** (3.96)	9.00** (0.0069)		77.4** (2.66)	151.9** (2.185)	6727	0.996
		Intercept	**HI**	**HI^2**	**EARLY**	**LATE**		
	NOx	539** (48.2)	0.171** (0.0245)	4.8E-05** (2.868E-06)	-496** (8.73)	13.8** (7.19)	6727	0.900

** indicates significance at the 5% level
* indicates significance at the 10% level

Figure 4-11 shows the observed, predicted, and counterfactual estimates of generation for the coal units in Plants 1571 and 1573. The predicted generation, shown by the line with marked with "x" is a reasonable estimate of the generation, although for the P1571U1, it does not quite fully predict the peaks and valleys. The counterfactual estimates of generation were slightly higher than the observed generation for both plants in the early ozone season and for P1573U2 during the entire ozone season. This is indicative of a reduction in output during the ozone season.[114] Figure 4-12 shows the counterfactual, predicted, and observed weekly NO_X emissions for the two generating

---

[114] Changes in electricity prices and fuel prices did account for some of the observed reductions in output during the ozone season because the counterfactual estimates of generation did decline at the beginning of the ozone season.

units. Both of these units abated NO_X emissions more during the beginning of the ozone season compared to the late ozone season.

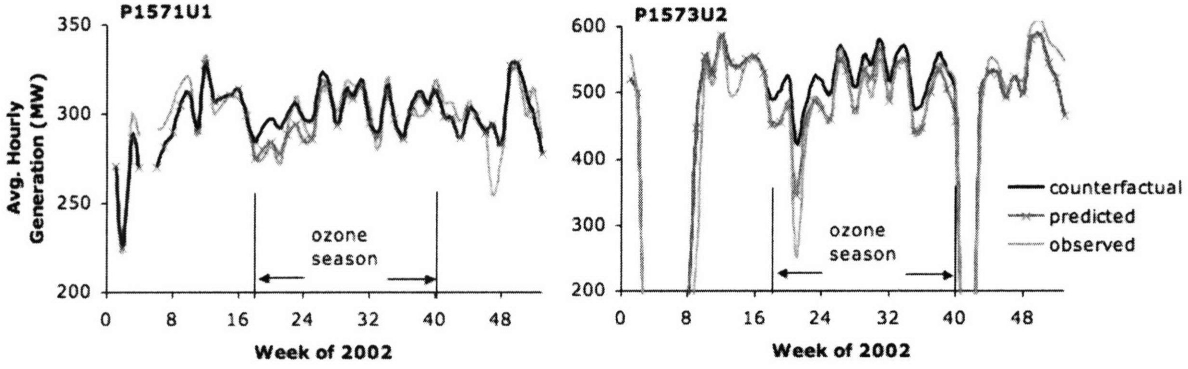

Figure 4-11 Observed, predicted, and counterfactual average hourly generation by week for two coal units in Plants 1571 and 1573 during 2002.

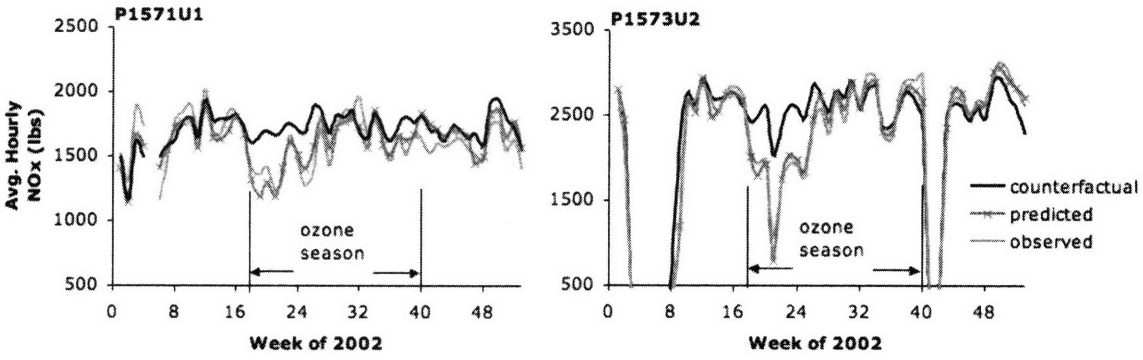

Figure 4-12 Observed, predicted, and counterfactual average hourly NO_X emissions by week for two coal units in Plants 1571 and 1573 during 2002.

About 70% of the NO_X abatement for P1571U1 occurred in May and June and nearly all (98%) occurred in May and June for P1573U2. Reductions in output caused about 87% of the NO_X abatement from P1571U1 in May and June and none of the observed abatement between July and September (calculated from the counterfactual emissions rate and the observed generation). Reductions in output caused about 43% of the NO_X abatement during May and June for P1573U2. The unit P1573U2 likely only used its control technology in May and June because it abated very little between July and September and reductions in output only explain 43% of the unit's abatement in May and June.

The counterfactual estimates provide evidence that the operators of coal fired power plants may have responded to incentives to reduce emissions (through reductions in output and the use of control technologies) more during periods when power prices were lower (Appendix A shows that power prices were lower in May and June in 2002 and 2005). The exact implication of this for the effectiveness of the seasonal NO_X cap and trade program is not easily determined because of the complicated chemistry of ozone formation. However, ozone episodes are less likely during May and June compared to the later months so this behavior is an example of how a differentiated program might provide a benefit. A successful differentiated cap-and-trade program would not require these reductions if they would not have helped some area achieve the ozone air quality standards.

The counterfactual estimates also provide examples of the tradeoff between heat rate and NO_X rate that was explained in Section 4.3.2. Table 4-4 summarizes the observed and counterfactual NO_X and heat rates for the two units for the early, late, and non-ozone seasons. Based on the discussion in Section 4.3.2 we would expect an increase in heat rate of about 3% to accompany a decrease in NO_X emission rate of about 30% from a generating unit employing low NO_X burners (which these did). The changes described in Table 4-4 were roughly this order of magnitude. The NO_X emission rate of P1573U1 decreased by about 19% and its heat rate increased by about 3.5%.

Table 4-4 Average NO_X emission rates and heat rates for two coal generating units in 2002. "Early" refers to the months of May and June, the early ozone season. "Late" refers to the months of July through September.

		NOx Rate			Heat Rate		
		Observed	CF	% Reduction	Observed	CF	% Increase
Plant 1571	**Early**	0.51	0.61	15	9.5	9.4	1.2
	Late	0.54	0.61	11	9.9	9.5	5.0
	Off	0.59	0.61	2.6	9.4	9.4	-0.2
Plant 1573	**Early**	0.42	0.52	19	9.7	9.4	3.5
	Late	0.53	0.53	-0.5	9.8	9.8	-0.6
	Off	0.52	0.53	0.4	9.5	9.3	3.1
		lbs/mmBTU		%	*mmBTU/MWh*		%

4.5.3. Incentive 3: NO_X control technologies reduce capacity

One of the implications of the increases in heat rate caused by the use of a NO_X control technology is a reduction in the generator's maximum output. As discussed in Section 4.4.3, generating units are limited by a heat input capacity; an increase in heat rate implies that the unit cannot generate as much from its maximum heat input.

As an example, the average heat rate of P1571U1 was 9.9 mmBTU/MWh during the late ozone season, compared to 9.5 mmBTU/MWh during the early ozone season of 2002 (Table 4-4). Although power prices were higher in the late ozone season, the unit's maximum generation was only 349 MW in the late ozone season, compared to 364 and 360 during the early and non-ozone seasons respectively. The unit did not generate above the annual 99th percentile of its hourly generation in the late ozone season.

The examples discussed in Sections 4.1 through 4.3 provide evidence that the operators of at least two power plants found it worthwhile to reduce output and NO_X emissions more in the early part of the ozone season, when power prices were low, than in the later parts of the ozone season. The observation of this behavior is consistent with the incentives discussed in Sections 3.1 through 3.3.

4.5.4. Aggregate effects of incentives for poorly timed NO_X reductions

This section presents counterfactual estimates of weekly NO_X emissions without the seasonal cap-and-trade programs in 2002 and 2005 for the coal generating units in OTC region of the PJM and New England power systems. The aggregate counterfactual was used to test hypothesis that more NO_X abatement occurred from these units during the early weeks of the ozone season (i.e. in May and June) compared to the later weeks. The estimates suggested that the weekly average NO_X reductions in May and June of 2002 were about 1930 tons, compared to 1510 tons in the months of July through September. In 2005, the weekly average NO_X reductions were about 3880 tons, compared to about 3380 in the later months. The estimates suggest that at least some power plant operators reduced emissions more in the earlier months, which is consistent with the incentives discussed in Section 4.4.1 through Section 4.4.3.

To construct the counterfactual estimates, weekly panel datasets were created for the years 1998, 2002 and 2005. The year 1998 was included a control because the seasonal NO_X cap-and-trade program first began in the summer of 1999. The datasets included observations on roughly 100 generating units for 52 weeks in each year. The dataset included weekly generation, heat input, and NO_X emissions for each unit. It also included weekly average electricity price (hub prices) and average weekly electricity demand for the New England and PJM power system. Only PJM and New England were included because electricity prices and total electricity demand data were not available for 1998 and 2002 for the New York Power Pool. Only the OTC region (and not the entire SIP Call region) was included so that the number of units in the analysis did not increase dramatically between 2002 and 2005.

The estimation procedures were similar to those for the counterfactual estimates for the individual units discussed in Section 4.5.2, but did not include a piecewise linear estimate of generation. The piecewise linear estimates of generation from electricity price

(Section 4.5.2) required high frequency data to determine the unique cut-point for each generating unit and this was not practical for the larger set of units. The aggregate estimates also used weekly average electricity demand to predict generation from the coal power plants rather than electricity prices. Primary reasons for this were that the average demand predicted generation better for all three years and this analysis did not necessitate estimating a coefficient for electricity price. In addition, total electricity demand is inelastic and therefore independent of the generation from any particular generator. As discussed below, the relationship between electricity demand and generation from the coal plants did not change significantly during the "ozone season" of 1998 (before the cap-and-trade program was implemented). This suggests that total electricity demand can be used as a reference point to determine whether or not the coal plants generated more or less as a result of the NO_X cap-and-trade programs in 2002 and 2005.

Instrumental variable regressions of output on electricity price (with total electricity demand as an instrumental variable) predicted generation reasonably well in 2002 but not in 1998 or 2005 (see Appendix C).[115] One possible reason that it was difficult to predict output from electricity prices consistently between the three years is that changes in the structure of wholesale electricity markets occurred between 1998 and 2005 because of liberalization and expansion (including in the middle of the year in 2005 when PJM added the Dominion Control Area). This could have changed the relationships between prices and generation during these years. In addition, electricity prices are disproportionately high (e.g. $1000/MWh) during scarcity conditions that occur when system operators cannot maintain operating reserves. Under these conditions, the prices are not set by the costs of the marginal generating unit – they are much higher. The

---

[115] Total weekly electricity demand was used as an instrument for average weekly electricity price in these regressions. Electricity prices are likely endogenous in the regression of generation on price because even if a coal unit is not marginal, the level of its generation could influence which other unit was marginal – and therefore the price. Total electricity demand is inelastic and therefore not determined by the generation of any particular coal unit. Electricity demand is also correlated with price.

regressions of generation on electricity price tended to over-predict generation during these price spikes (the piecewise linear estimates were able to address this problem for the estimates of generation for individual units in Section 4.4.2).

For the aggregate estimates, ordinary least squares estimates using the fixed effects model were used to create counterfactuals for generation, heat input, and NO_X emissions. To predict generation, the coefficients in the following equation were estimated:

$$q_{it} = \beta_o + \beta_{1i}\mathrm{DEMAND}_t + \beta_{3i}\mathrm{EARLY} + \beta_{4i}\mathrm{LATE} + \varepsilon. \qquad (11)$$

where DEMAND was the weekly average electricity demand in the power system corresponding to each generating unit (i.e. PJM or New England), EARLY was a dummy variables for May and June and LATE for July through August. Weekly heat input was then regressed on the predicted generation ($\tilde{q}_{it}$) from Equation 11 and on the dummy variables:

$$HI_{it} = \alpha_o + \alpha_{1i}\tilde{q}_{it} + \alpha_{2i}\mathrm{EARLY} + \alpha_{3i}\mathrm{LATE} + \varepsilon \qquad (12)$$

In the last stage, weekly NO_X emissions were regressed on the predicted heat input ($\tilde{HI}_{it}$) from Equation 12, the square of the predicted heat input and the dummy variables:

$$NO_{Xit} = \mu_o + \mu_{1i}\tilde{HI}_{it} + \mu_{2i}(\tilde{HI}_{it})^2 + \mu_{3i}\mathrm{EARLY} + \alpha_{4i}\mathrm{LATE} + \varepsilon \qquad (13).$$

The "staged" regressions (regressing the heat input on the predicted generation in Equation 12 and the NO_X emissions on the predicted heat input in Equation 13) were performed because generation, heat input and NO_X emissions are simultaneously determined in this system. For example, the NO_X emissions from a given level of heat input are determined by whether or not a NO_X control strategy was used and NO_X control strategies can affect a generating units' efficiency (how much heat input is required to generate a given amount). The staged regressions insured that only the variation in generation due to changes in total electricity demand was used to explain variations in heat input and, in turn, that only the variations in heat input due to electricity demand (and not NO_X abatement decisions) were used to predict NO_X emissions.

Table 4-5 shows the coefficients estimated in these panel regressions. The coefficients on the EARLY and LATE dummy variables were negative and significant in the 2002 and 2005 regressions of generation on total electricity demand (except the coefficient on LATE in the 2005 regression was not significant). The coefficients on the dummy variables were not significant in 1998. This suggests that the seasonal NO_X cap-and-trade programs caused a reduction in output from this set of coal generating units. In addition, the hypothesis that the coefficients on the LATE and EARLY dummy variables were equal was rejected at a 5% level of significance in the 2002 and 2005 regressions of generation on demand. This indicated that a larger reduction in generation occurred in May and June compared to July through August.

The coefficients on the dummy variables in the regressions of NO_X emissions on heat input were negative and significant in all years, but were larger in magnitude in 2002 and 2005. The significant coefficients on the dummy variables in 1998 suggest that slight NO_X reductions for a given level of heat input occur during the summer even without the NO_X cap-and-trade programs. However, the coefficient on EARLY was smaller in magnitude than that on LATE in 1998 (although this difference was not significant at the 5% level). The opposite was true in 2002 and 2005. This suggests that more NO_X abatement occurred in May and June in 2002 and 2005 than in July through August, which is consistent with the incentives for more NO_X abatement in the early summer. However, in both years, the hypothesis that the coefficients on EARLY and LATE were equal could not be rejected.

The estimated coefficients in Equations 11 through 13 for the non-ozone season (i.e. without the ozone season dummy variables) were used to predict the counterfactual generation, heat input, and NO_X emissions. Figure 4-13 shows the observed, predicted, and counterfactual NO_X emissions for the three years. The graphs show that, especially in 1998 and 2002, the regressions predicted the non-ozone season NO_X emissions reasonably well. Figure 4-14 shows that the estimated weekly NO_X reductions were higher in the early weeks of the 2002 and 2005 ozone seasons and declined throughout the ozone season. In 2002, the average weekly NO_X reductions were larger in May and June, compared to July through August, by about 430 tons. In 2005 this difference was about 500 tons. The observation of addition reductions early in the summer is consistent with the incentives discussed in for power plant operators to reduce NO_X emissions more when power prices were lower and ozone formation less likely (in the early ozone season). However, the estimation procedure also suggested that slightly more NO_X reductions in 1998 – which is used as a control since it was the final year without a seasonal cap-and-trade program – occurred in May and June. The difference between the average reductions in May and June compared to July through August in 1998 was about 130 tons. This suggests that mechanisms not related to the seasonal cap-and-trade programs may have caused a portion of the observed "extra" NO_X reductions in May and June in 2002 and 2005.

Table 4-5 Predicted coefficients from regressions of generation on total electricity demand, heat input on predicted generation, and NO_X emissions on predicted heat input in 1998, 2002, and 2005 using weekly panel data for coal plants in the OTC region in the New England and PJM power systems.

1998	Regression of Generation on Total Demand		Regression of Heat Input on Predicted Gen		Regression of NOx Emissons on Predicted Heat Input	
Coefficient	**Generation (gen)**	**Standard Error**	**Heat Input (HI)**	**Standard Error**	**NOx**	**Standard Error**
Demand	1.74**	0.15	--	--	--	--
Early	-311	838	4171	8040	-3.8*	2.1
Late	59	952	4036	9218	-5.3**	2.5
Constant	-11770**	3777	-4354	26763	-45.9*	25.5
Predicted Gen	--	--	9.91**	0.83	--	--
Predicted HI	--	--	--	--	0.00049**	0.00014
Predicted HI^2	--	--	--	--	-2.9E-10*	1.69E-10
Number of Observations		4440				
Number of Groups		88				

2002						
Coefficient	**Generation (gen)**	**Standard Error**	**Heat Input (HI)**	**Standard Error**	**NOx**	**Standard Error**
Demand	0.65**	0.07	--	--	--	--
Early	-3117**	620	20890**	4815	-11.3**	1.1
Late	-1668**	650	21922**	4239	-11.0**	1.2
Constant	6687**	2062	-3124	21139	-78.7**	24.9
Predicted Gen	--	--	9.75**	0.84	--	--
Predicted HI	--	--	--	--	0.00083**	0.00018
Predicted HI^2	--	--	--	--	-1.17E-9**	3.23E-10
Number of Observations		6869				
Number of Groups		137				

2005						
Coefficient	**Generation (gen)**	**Standard Error**	**Heat Input (HI)**	**Standard Error**	**NOx**	**Standard Error**
Demand	0.50**	0.06	--	--	--	--
Early	-1140**	568	2880	4057	-31.7**	1.47
Late	-215	580	6537*	3982	-30.3**	1.59
Constant	35953**	1774	11436	42310	--	--
Predicted Gen	--	--	6.58**	0.83	-909**	167
Predicted HI	--	--	--	--	0.0053**	0.0009
Predicted HI^2	--	--	--	--	-7.24E-9**	1.28E-09
Number of Observations		4840				
Number of Groups		104				

** indicates significance at the 5% level
* indicates significance at the 10% level

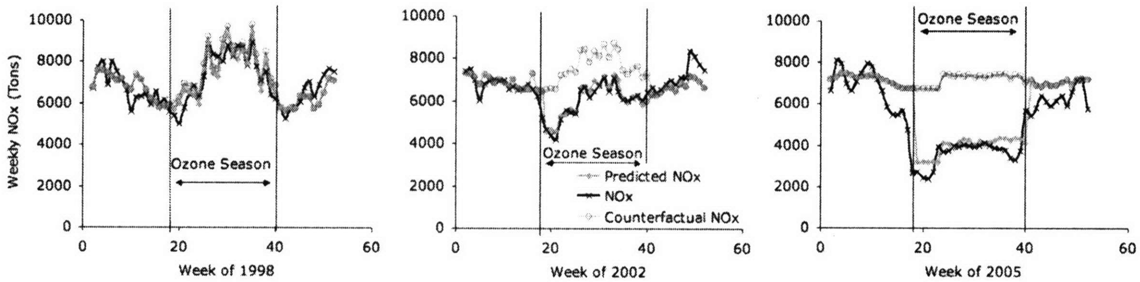

Figure 4-13 Observed, predicted, and counterfactual NO_X emissions for coal generating units in the OTC region in the PJM and New England power systems in 1998, 2002, and 2005.

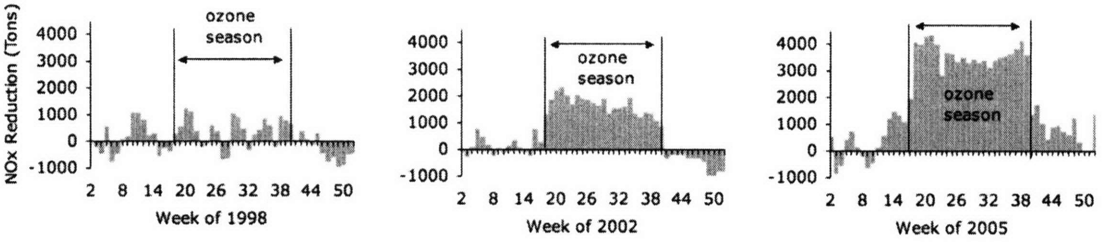

Figure 4-14 Weekly NO_X reductions in 2002 (middle) and 2005 (right) for coal plants in New England and PJM.

4.6. Summary

The problem of ozone pollution creates a challenge for environmental regulation. The effectiveness of NO_X reductions in mitigating peak ozone concentrations depends on timing and location and fluctuates with meteorological changes. Seasonal cap-and-trade programs have caused large stationary sources, like coal power plants, to reduce NO_X emissions in the Eastern U.S. during the summer since 1999. Air quality monitoring data suggest that these reductions have helped reduce the frequency of high ozone concentrations in many areas, but not all highly populated areas in the Eastern United States have attained the ozone standard.

A prescriptive regulation to address ozone nonattainment might mandate specific emission rates or control technologies for particular power plants – such as those that were uncontrolled or those that operated more during hours with high electricity demand and prices.[116] But the heterogeneity in the NO_X emission characteristics of coal power plants alone, not to mention natural gas and oil generating units, suggests that it would be difficult for regulators to determine the most cost-effective NO_X control options, making

---

[116] Recent regulations in some OTC states are requiring NO_X reductions from peaking units. But, neither the EPA nor state regulators have performed air quality modeling that indicates that the reductions will affect peak ozone concentrations in non-attainment areas (personal conversations with EPA and state regulators). See the OTC's "Memorandum of Understanding Among the States of the Ozone Transport Commission Concerning the Incorporation of High Electrical Demand Day Emission Reduction Strategies into Ozone Attainment State Implementation Planning," March 2, 2007

decentralized abatement decisions under a market-based regulation preferable. In addition, the prescriptive measures would not guarantee that the reductions would mitigate ozone formation, especially since the places and times in which reductions are needed vary with the weather.

When considering regulatory options for obtaining further NO_X emission reductions from large stationary sources, reductions in the seasonal cap of the current NO_X cap-and-trade program could also be an inefficient, and potentially ineffective, solution. Seasonal cap-and-trade programs to not guarantee that only the most cost-effective NO_X emission reductions occur. Indeed, under these programs the operators of coal power plants have incentives to reduce NO_X emissions when electricity prices are low, which are also times when ozone formation is least likely. Three incentives were discussed that suggested that it could be less costly for the operators of some coal generating units to abate NO_X emissions when power prices were low. First, output reduction is an effective NO_X control strategy because some coal units' NO_X emission rates decrease at decreased levels of output and reduced output is less costly when power prices are low. Second, if a generating unit's marginal costs increase more steeply with a NO_X control technology than without it the operator might prefer to use the technology only during periods with low expected power prices. Third, if a NO_X control technology increases a generating unit's heat rate it can decrease its effective capacity, which is more costly when power prices are high.

The estimated NO_X abatement of example power plants under the OTC NO_X Budget Program in 2002 was consistent with these incentives. In one example, nearly all of a generating unit's NO_X abatement occurred in the first two months of the ozone season when power prices were lower than in the latter three months. Aggregate estimates of NO_X abatement in 2002 and 2005 from coal generating units in the OTC regions of the New England and PJM power systems suggested that, although NO_X

reductions occurred during the entire ozone season, slightly more abatement took place in the earlier weeks. Atmospheric chemistry modeling would be needed to determine whether the additional NO_X reductions of about 400 tons per week that occurred in the early weeks of the ozone season might have helped mitigate peak ozone concentrations if they had also occurred in the later weeks. But this chapter shows that a differentiated cap-and-trade program could improve the cost-effectiveness of the current cap-and-trade programs by providing stronger incentives for the emission reductions with the greatest potential to reduce ozone concentrations.

Chapter 5 – Flexible NO_X Reductions through Power Plant Redispatch in Classic PJM[117]

The literature has called for a more finely differentiated regulation of NO_X emissions to address the temporal and locational variation in the contribution of NO_X to ozone formation and associated damages to human health and welfare (Chapters 2 and 3). The examination of the NO_X characteristics of coal power plants supports this need (Chapter 4). This chapter examines one of the necessary conditions to apply such a program to stationary sources: that electric generators would have sufficient flexibility to reduce NO_X emissions in the short term if faced with higher NO_X prices on summer days when ozone formation is a problem.

One method to obtain short-term NO_X emission reductions from power plants is through redispatch: substituting generation from units with high NO_X emission rates with generation from the units with the lowest NO_X emission rates. Full utilization of capacity and network constraints could limit opportunities to substitute to low-NO_X emitting generators. Because electricity demand is inelastic, if substitutions (or "redispatch") were not possible, a higher NO_X price would increase electricity prices instead of achieving reductions in NO_X. In addition, there is the concern that reductions in one or several areas would create "hot spots" or higher NO_X emissions in another area.

The results of the simulations described in this chapter suggest that flexibility to reduce emissions from power plants through redispatch does exist, even in high demand hours. The capability to reduce NO_X through the substitution of generators exists because

---

[117] This work also appears in the following working paper: Martin, K.C., P.L. Joskow, and A.D. Ellerman (2007), "Time and Location Differentiated NO_X Control in Competitive Electricity Markets Using Cap-and-Trade Mechanisms," *Center for Energy and Environmental Policy Research Working Paper.*

of considerable variation in commitment among generators, even in high demand periods, and heterogeneity in emission rates, including that for coal plants described in Chapter 4. Heterogeneity in emission rates is often overlooked because models of NO_X emissions from power plants typically use emission rates that aggregate over region, month, and rarely by time of day or in response to specific operating conditions. This type of aggregation does not capture the full range of variation in power plant utilization and in heterogeneity of emission rates.

Moreover, as discussed in Chapters 2 and 3, the complex relationship between NO_X emissions, temperature, and atmospheric chemistry means that reductions in nighttime NO_X emissions could do more to mitigate peak ozone concentrations than reductions in daytime emissions. The potential to reduce NO_X emissions through redispatch increases at lower levels of demand, such as those observed at night during the summer. While this chapter considers only the potential reductions from redispatch, the analysis in Chapter 4 suggested that operators have some ability to control NO_X emissions rates of generating units through control equipment, changes in boiler combustion attributes, and through fuel switching. In the longer run, time and locational differentiated NO_X prices could affect investments in NO_X control equipment, boiler and turbine equipment.

5.1. Methodology

Two complementary methods were used to simulate the potential magnitude of reductions in NO_X emissions that can be achieved as a consequence of redispatch while meeting electricity demand and transmission network constraints in the "Classic" PJM power system. The "Classic" PJM power system is the original Pennsylvania, New Jersey, Maryland power system that included these three states as well as the District of Columbia and Delaware. PJM has now expanded, this original areas is generally upwind of the remaining ozone nonattainment areas (U.S. EPA 2006b). Both methods used generator-level emission rates to simulate Classic PJM and balanced electricity supply and demand. The "zonal" method accurately incorporated emission rates and historical load characteristics to demonstrate the physical potential for significant NO_X reductions through redispatch. Optimal power flow (OPF) and Security Constraint OPF (SCOPF) estimated both the physical feasibility of redispatching generators to reduce NO_X emissions and the levels of NO_X permit prices required to induce economic redispatch through wholesale market mechanisms.[118] The OPF and SCOPF simulations used PowerWorld Simulator and modeled network constraints more accurately than the zonal model. The two methods produced reasonably consistent results.

5.1.1. Background: Electric Power Systems and Wholesale Electricity Markets

The process of electricity production and delivery includes the generation of power as well as its transmission and distribution and the provision of "ancillary" services related to reliability.[119] The power plants in the Eastern U.S. that participate in seasonal NO_X cap-and-trade programs provide bulk power (generation) and ancillary services. The transmission network is also important for these generators because the capability of the network to transmit power over long distances determines which of the generators in the wholesale power market can be used to fill demand. The transmission network therefore partially determines both the costs and the environmental impacts of the generation used to fill electricity demand.

---

[118] The independent market monitor for PJM does not believe market power to be a significant problem in PJM, see PJM (2006) pages 59-69 and 83-93. For this reason, the NO_X price simulations assume that generating units engaged in Bertrand competition and bid their marginal costs into the PJM markets. The capabilities of PowerWorld allow exploration of the implications of market power and this is an opportunity for future research. For examples of work on the interactions of market power and emissions in PJM see Mansur 2006a and 2006b. Mansur (2006a) found that the exercise of market power in the PJM region leads to lower emissions and that, in this situation, a tradable permit system is superior to a tax in terms of welfare effects. Mansur (2006b) also found that electricity restructuring and the accompanying exercises of market power explained about one third of the emissions reductions observed when PJM restructured in 1999 and when the NO_X cap-and-trade program first took effect in the ozone transport region.

[119] Ancillary services include spinning reserves, frequency control, voltage support, and black start services. For a more complete description than that provided here, see generally Stoft 2005.

Most power plants in the Eastern U.S. that participate in the seasonal NO_X programs also participate in one of three deregulated, or competitive, wholesale electricity markets: the New England Power Pool, the New York Power Pool, or the PJM Interconnect.[120] Although these power systems have unique characteristics, their basic structures are similar and are important for analysis of past and potential behavior under NO_X cap-and-trade programs. All of the systems have "system operators" that coordinate the balancing of supply and demand and the provision of ancillary services. The real-time and day-ahead wholesale electricity markets in these regions use security constrained bid-based dispatch auction mechanisms that yield locational prices for electricity (e.g. Joskow 2006). The remainder of this subsection provides a brief explanation of these concepts (for further detail see Hogan 1998, Stoft 2002, and Joskow 2006).

Bid-based economic dispatch refers to a system in which generating unit owners bid cost curves to the system operator, who uses these curves to decide which generating units to "dispatch" to fill demand while minimizing total system operating costs subject to network and security constraints. A bid defines the minimum price at which a generator will provide a quantity of power on an hourly basis. If the network constraints allow, this process calls on generating units to provide power in "merit order": those with the lowest bids are used first to fill demand. In a competitive market the bids reflect the generating units' marginal costs of increasing output. The price of electricity is the cost to supply the system with one more megawatt of power; in a competitive market this is the marginal cost of the marginal generator – the generator that would be used to fill the next increment of demand. A cap-and-trade program that places a price on emissions increases a generating unit's costs and, under bid-based economic dispatch, the unit operator's bids will presumably reflect the additional costs associated with emissions when a cap-and-trade program is in place.

---

[120] The latter was formerly the "Pennsylvania, New Jersey, Maryland" power system but has expanded considerably and is now referred to only as "PJM".

A number of complications to the straightforward concept of economic dispatch are important affect the potential short-term responses of power plants to a high NO_X price like that from a differentiated cap-and-trade program. For any given hour, economic dispatch to meet electricity demand on a power network results in the transfer of electricity between network nodes according to complex but well understood physical laws. Thermal transmission constraints that limit the amount of power than can flow across transmission equipment can prevent the system operator from implementing the lowest cost dispatch. A transmission line, or other piece of equipment like a transformer, becomes "congested" if the thermal limitations on the amount of power it can carry cause generators to be dispatched out of merit order. In this case, a lower cost generator could have provided more power if the transmission equipment was not constraining.

The related concept of locational marginal pricing assigns prices to the nodes of the power network that reflect the costs imparted by congestion and losses as well as the marginal costs of generation.[121] On a power network with no transmission constraints and no physical losses, the generator with the lowest marginal cost can always fill an increment of demand at any node and, as a result, all nodes on the network have the same price for electricity. However, network congestion means that the lowest cost generator cannot be used to fill an increment of demand at some nodes. The locational marginal price of "import-constrained" nodes is higher to reflect the higher marginal cost of generators needed to fill any additional demand at the nodes. Thus, prices at different nodes on the network vary to account for the marginal cost of congestion and the marginal cost of losses. The marginal costs of losses vary across the system because losses themselves vary with the particular impedances of lines and transformers.

---

[121] The wholesale electricity spot markets in New England and New York have included the marginal cost of losses in locational prices for a number of years. The PJM Interconnection, which we focus on here, began including the marginal cost of losses in its locational pricing mechanism in 2007 – after the 2005 summertime period studied here.

System operators also consider security constraints in addition to physical network constraints. A single contingency (e.g. line or power plant outage) can cause cascading failures on a power network. To prevent catastrophic failure, system operators use security-constrained dispatch: they operate the system to minimize the potential impact of any major contingencies. For example, a contingency might be that the line from A to B opens unexpectedly. If the line from A to B opens, the flows across the other lines in the network change and the event could cause overflows or faults on other equipment. A security-constrained economic dispatch that considered the contingency of the line from A to B opening might minimize operating costs subject to the constraint that if the line from A to B opened, no other overflows would occur. The system operators consider sets of hundreds to thousands of contingencies.[122]

The redispatch simulations performed to estimate the flexibility to reduce NO_X emissions assume that if generating units' costs change so too will their bids and, if the new costs change the merit-order of the generating units, the resulting security-constrained economic dispatch will change as well. A new dispatch of generating units will likely create a new set the binding network constraints (both thermal and contingency). The new patterns of congestion and the increases in generators' marginal costs due to the NO_X price will be reflected in the locational marginal prices (LMP).

5.1.2. *Simulating NO_X Reductions from the Redispatch of Power Plants*

The redispatch simulations were designed to account for the interactions between generator marginal costs, congestion, and dispatch. The simplest simulations used a zonal model to identify portions of the Classic PJM network that were reasonable approximations of areas where the transmission system was capable of handling the exchange of generation between units without causing congestion on the lines between

---

[122] The PJM system operator, the largest of the three power systems, considers over 1000 contingencies: PJM, "FTR Model Information," at http://www.pjm.com/markets/ftr/model-info.html.

zones or severely altering flows on these lines. That is, the analysis identified interconnected zones of generating units that could be considered good physical substitutes for each other during the ozone season of 2005. Substitution between zones was assumed to be infeasible if it required increasing power flows from one zone to another zone where network constraints were binding.

There has been a debate in the academic literature over the relative merits of zonal and nodal pricing systems (e.g. Stoft 1997 and Hogan 1999). The literature shows that the complexities caused by flows over parallel lines in electricity networks, and the variations in those flows over time due to fluctuating demand, make it difficult to create consistent zones by collecting nodes that have the same or similar LMPs (Stoft 1997). But the zonal model captured many of the details of the Classic PJM power system important for estimating potential NO_X reductions – like the actual emission rates of generating units in PJM and the locations of generation and of congested lines – while using only publicly available data and a relatively simple characterization of the topology of the transmission network.

PowerWorld's OPF and SCOPF capabilities were used to capture a richer characterization of network power flows and constraints. The model used power flow bases cases parameterized to match the classical PJM network as a second method to estimate the physical capabilities to reduce NO_X emissions. This model provided a more refined account of the physical complexities, constraints, contingencies and parallel flows on the network. However, despite this method's ability to model nodal prices, the parameters of the network used in this model also change in time due to fluctuating demand. Any feasible representation of an electric power system will not capture how its electrical properties change in real time with patterns and levels of utilization and with ambient conditions.

5.1.3. Construction of the Zonal Model

Publicly available data on the PJM transmission system[123] – on the name, type (e.g. generator, load), and voltage of each bus and the buses to which each connects – were used to create an abstract representation of the PJM system, or a network graph.[124] The network graph represents the substations, as nodes, and the inter-substation transmission lines, as arcs, between the nodes. Substations were defined broadly as closely connected collections of electrical equipment. Examples are a power plant with multiple generators and transformers, multiple power plants, or a switching station.

The data were matched by substation name into a system that includes over 900 nodes and over 8500 connecting lines. The substation names, voltages, and equipment were used to match the generators in the EPA's Continuous Emissions Monitoring System (CEMS) to the nodes.[125] Hourly generating unit operation data, like heat input, generation, and emissions, are available from the CEMS data. These data are available for fossil fuel-fired generating units with rated capacities of at least 15 or 25 MW, depending on the state. The same EPA website houses data on the characteristics of emission sources like their location, technology type (e.g. dry bottom wall-fired boiler), types of fuel burned, the sources' emission control technologies, and when they installed these control technologies. Less detailed data on the rated capacities of other types of generating units (e.g. nuclear, hydro, and municipal waste) and smaller units are available from the Energy Information Association (EIA).[126]

---

[123] PJM, "Transmission Facilities," available at http://www.pjm.com/services/transm-facilities.jsp.

[124] Network graphs are used in the mathematical field of graph theory, computer science, and social network theory. They are abstractions that model pairwise relationships between objects using nodes (e.g substations) and "edges", "arcs", or "lines" (in this case transmission lines). For other applications of network theory to electric power systems see Watts (1998).

[125] See Environmental Protection Agency's Continuous Emissions Monitoring System (CEMS) (unit generation and heat input data) and data on emissions and characteristics of regulated sources at http://cfpub.epa.gov/gdm/.

[126] See EIA "Form EIA-860 Database: Annual Electric Generator Report," available at http://www.eia.doe.gov/cneaf/electricity/page/eia860.html.

Using these publicly available data sources, approximately 49.1 GW of fossil fuel-fired capacity (rated summertime capacity) in the EIA's database of existing capacity were matched to the appropriate substation in the PJM network graph. The 2005 PJM State of the Market Report states that there were about 50.6 GW of fossil capacity in PJM in 2005 (PJM 2006); the matching process covers about 97% of the fossil capacity in PJM. Of the 49.1 GW capacity in the EIA database, about 96% of it (47.2 GW) reports emission data to the EPA's CEMS database. In all this yielded detailed data on the emissions from about 93% of the fossil fuel-fired capacity in PJM.

Two criteria were used to create zones in the PJM network within which congestion rarely occurred. In its State of the Market Report, PJM discusses the impact of frequently congested lines on market concentration (PJM 2006). For 2005, it lists thirteen transmission lines and transformers that were congested for over 100 hours in 2005. In addition, the State of the Market Report discusses three other lines and one other transformer that were frequently congested in 2004. The first criterion used to identify zones within PJM was that these 17 lines must be located on the borders *between* zones and not within the zones.

The second criterion used historical hourly locational marginal price (LMP) data to define zones and created smaller zones than the first criterion alone. Hourly LMP and zonal demand data for PJM are available on the PJM website and were matched to the network graph.[127] The second criterion was that the standard deviation of the LMP's within each zone was less than $10/MWh for at least 90% of a sample of 144 summertime hours in 2005. This criterion was selected because differences in LMP of less than $10/MWh rarely indicate congestion; more typically, they indicate other

---

[127] PJM website "Real Time" energy market data at http://www.pjm.com/markets/energy-market/real-time.html.

differences in marginal cost between nodes.[128] Additionally, the zonal model was designed to capture only the most frequent patterns of congestion, not every pattern that occurred. Many of the identified zones easily met the LMP criterion. For example, in the largest zone of 117 nodes, the standard deviation of LMPs was less than $5/MWh in 90% of the hours and less than $10/MWh in 98% of hours.

These two criteria created 35 zones with between 117 and 4 nodes in each. The network graph was then used to match generating unit emissions and generation data to the zones. This matching allowed estimation of the potential reductions in NO_X from redispatch while taking account of the constraints caused by the most frequent patterns of network congestion in 2005. To estimate the maximum potential NO_X reductions, or the technical upper bound on NO_X reductions from redispatch, NO_X emissions from the fossil fuel-fired generating units in Classic PJM were minimized subject to five constraints for each hour of analysis:

1. Total generation from the generators was held constant.
2. The generation from any unit operating in the hour could only be reduced to 20% of its rated capacity; not to zero.
3. Only combustion turbine units could "turn on", all other units remained off if they did not generate in the hour. Generating units could produce power up to 100% of rated summer capacity.[129]
4. The total generation from all the generating units in zones on the high-LMP side of congested lines and transformers could not decrease.
5. The total generation from all the generating units in zones on the low-LMP sides of congested lines and transformers could not increase.

---

[128] PJM lists both LMPs and data on "real time constraints" or "transmission limits". The LMPs between nodes often vary up to $30/MWh without the line between those nodes being listed as a constraint. See "PJM Operational Data" at http://www.pjm.com/pub/account/lmpgen/lmppost.html or "Real Time Transmission Constraints 1998-2005" at http://www.pjm.com/markets/energy-market/real-time.html.

[129] The summer rated capacities used in these simulations do not reflect forced outages. In PJM for 2005, the demand equivalent forced outage rate was about 7.3% (PJM 2006). The forced outage rates are not available for the summer months when high electricity prices provide an extra incentive for plants to be available and operational. The fact that some plants might not be able to turn on or up to 100% of their capacities makes estimates that do not account for outage rates optimistic. We include some simulations that account for annual average forced outage rates. These estimates are slightly restrictive because summertime forced outage rates tend to be lower than the annual rates.

The second of these constraints reflected the high start-up costs that prevent some generators from being turned off often and it also maintained levels of operating reserves. The third reflects the time and costs associated with start-up for units other than combustion turbines, which can start quickly. The fourth constraint was necessary because a decrease in the net generation from units on the high-LMP side of a constraint would cause an increase in the power flowing over a congested line. Similarly, the fifth constraint was necessary because increasing the net share of power from units on the low-LMP side of a constraint would necessitate an increase in the power flowing over the congested line. It is possible, however, to increase the generation from units on the high-LMP side of a congested line while reducing that from the generators on the low-LMP side. This would decrease the flow of power over that line (i.e. create counterflow), thereby relieving congestion.

An additional assumption used in this analysis was that the NO_X rates for the units generating electricity in a given hour did not change from those observed in that hour, regardless of any changes in the quantity generated by that unit or changes in the utilization of NO_X control equipment or changes in combustion attributes. If the unit was not initially operating in an hour, its NO_X emissions were estimated based on its average NO_X rate for the hours between May 1st and September 30th, 2005. This assumption is likely to underestimate the potential NO_X reductions because the emission rates of many coal units decrease with decreasing utilization and generators have some flexibility to vary NO_X emissions rates in the short run (Chapter 4).

The zonal model was used to estimate NO_X reductions for a 24-hour diurnal period between August 3rd, 2005 at 2pm and August 4th, 2005 at 2pm as well as for various other hours during the summer of 2005. Three variations of the analysis also tested the impact of the five constraints listed above on the results. First, relaxing the fourth and fifth constraints estimated the "unconstrained" case: the potential NO_X reductions if network constraints were not a factor. In the second variation, the capacities of generating units were derated by the forced outage rate for PJM in 2005 (about 7.3%, see PJM 2006). Last, the most restricted case strengthened the third constraint by redispatching only the unused ("excess") capacity of generating units that were already operating in each hour to estimate potential NO_X reductions (i.e. no generating units were "turned *on*").

The zonal analysis has three major limitations. First, it does not consider new network overloads that the redispatch of generating units might cause. Second, it does not consider the loop flows at the borders of zones that might require units on the either side of a constraint to increase or decrease their output in order to avoid an increase in the flow over a congested line. Third, it does not consider contingency constraints. The second method using a security constrained optimal power flow (SCOPF) model of the PJM network, described immediately below, helps to address these issues.

5.1.4. *Security Constrained Optimal Power Flow with PowerWorld Simulator®*

PowerWorld Simulator contains a security constrained optimal power flow (SCOPF) analysis package that can solve power flows for large electricity systems while optimally dispatching generators and enforcing transmission limits, interface limits, and contingency constraints.[130] PowerWorld was used to simulate how a range of uniform NO_X permit prices for Classic PJM, incorporated into linear cost curves for generators, changed the security constrained economic dispatch of those generators. This exercise

---

[130] PowerWorld uses a full Newton-Raphson AC load flow algorithm or a DC approximation to solve the power flow. The optimal power flow capability simulates economic dispatch by iterating between solving the power flow and minimizing total system operating cost, using generator cost-curves, while enforcing system constraints like line and generator operating limits. Thus, the security constrained optimal power flow simulates economic dispatch while enforcing both normal operating limits and ensuring that there are no operating limit violations during specified contingencies (PowerWorld Corporation at http://www.powerworld.com/). For more explanation of the widely used algorithms behind optimal power flow models such as PowerWorld Simulator see, for example, Sun *et. al.* (1984).

also estimated the NO_X prices needed to achieve a range of NO_X reductions up to the maximum level (i.e. when further increases in NO_X prices caused little additional reductions). This analysis provided a measure of the physical capability to alter NO_X emissions from redispatch, estimated the NO_X prices required to induce different levels of NO_X emissions through redispatch of generating units.

"Base-case" or "solved" power flow models are one way to convey information about the network elements in a power system. Base-case power flows include data like the voltages and impedances of the elements in a power network as well as characteristics of power plants and loads. The information can either be very detailed or highly aggregated. The base-case power flows also include predetermined power injections at generator nodes and power withdrawals at load nodes. Base-case power flows are typically specified for a season or month because the characteristics of equipment change with ambient temperature. The voltages and impedances, and the implied physical limits, of the network equipment represent the network for the specified load level and ambient conditions. This information allows a program like PowerWorld simulator to solve for the power flows across the network.

Adding generator cost and capacity information to a base-base power flow enables optimal power flow simulations, which minimized total cost of operating the power system subject to network and security constraints. In the Classic PJM simulations, the variable costs of the power plants were represented by linear cost curves (i.e. constant marginal cost curves).[131] The linear cost curves were defined simply by incorporating NO_X emissions as an addition fuel cost:

$$c_i\,(\$/MWh) = H_i(p_{fi} + p_{ni}N_i) + O\&M_i$$

---

[131] The generation and load in areas of PJM outside the Classic PJM footprint were held constant between the base case and the "redispatched" cases. The generation and load in the areas surrounding the larger PJM were zero in the base case and subsequent cases; thus imports and exports to and from PJM as a whole were assumed to be zero.

where, for each generating unit *i*, H_i is its heat rate (mmBTU/MWh), p_{fi} is the price of fuel ($/mmBTU), p_{ni} is the price of NO_X permits ($/ton), N_i is the unit's NO_X emission rate in (tons/mmBTU), and $O\&M_i$ is the unit's variable O&M costs in ($/MWh). For each level of demand and NO_X price, the units were "dispatched" in order of least cost according to these cost curves. The NO_X price was applied uniformly to all units in PJM and was varied between $2000/ton and $125,000/ton.[132]

Data on the average delivered cost of fuel for natural gas, coal, petroleum products, and petroleum coke delivered to the electricity sector from the EIA's Electric Power Monthly for August 2005 were used to generate the cost curves. These data were matched to the generating units by state and fuel. The variable O&M data were from the Annual Energy Outlook for 2006 matched roughly by technology type and fuel.[133] The EPA CEMS provided data on 2005 ozone-season heat rates and NO_X emission rates.

As in the zonal model, we compared the NO_X emissions resulting from three cases: 1) an "unconstrained" case where the generation from units in Classic PJM was dispatched economically without enforcing network constraints, 2) the constrained case (optimal power flow "OPF") in which the network constraints, like line limits, were enforced, and 3) the security constrained case in which both network and contingency limits were enforced (security constrained optimal power flow "SCOPF"). In this way, the PowerWorld analysis complemented the zonal analysis, which did not address security constraints or whether redispatch created new congestion. To mimic the zonal analysis, we designated only combustion turbine units as "fast start" generators. This meant that the dispatch algorithms could turn on combustion turbines, but could only increase or decrease the output of all other units. As in the zonal analysis, we constrained the generation from all initially operating units to be at least 20% of their capacity and

---

[132] In August of 2005 these prices were around $2500/ton. Prices are currently about $1000/ton.

[133] (U.S. EIA 2006b) Table 38, page 77 and EIA's *Electric Power Monthly*, Tables 4-10 through 4-13, available at http://www.eia.doe.gov/cneaf/electricity/epm/epm_ex_bkis.html.

units could generate up to 100% of their summertime rated capacities. We also held the generation from all units outside Classic PJM and imports and exports constant.

Two base-case power flow models were obtained to simulate Classic PJM and each contained different data. The PJM Financial Transmission Rights (FTR) base-case power flows contain nodal loads and power injections for average levels of demand for hours in different months.[134] They do not include detailed information about generator capacities or NO_X rates, only that the generators existed at certain buses and that some produced a given amount of power in the modeled hour. Also, the generating unit identifiers in the PJM FTR model and the EPA and EIA capacity and NO_X rate data are not the same so matching the EPA and EIA data with the correct buses in the FTR model was a challenge and required some assumptions (see Appendix D). The model information for the FTR power flows includes a list of contingency (or security) constraints that PJM considers.[135]

The FTR base case power flows simulated hours with average electricity demand, around 38 GW in Classic PJM, which was typically in Classic PJM during nighttime hours in hottest parts of the summer.[136] NO_X reductions in nighttime hours may be important for ozone formation because, for example, winds can transport nighttime NO_X emissions to highly populated areas where ozone can form during the day. The integration of this work with atmospheric chemistry models will eventually show whether NO_X emission reductions during nighttime or daytime hours will most effectively reduce ozone concentrations.

---

[134] The PJM FTR base-case power flows are available at PJM, "FTR Model Information," http://www.pjm.com/markets/ftr/model-info.html. These cases are available publicly to participants in the PJM wholesale markets but require a password, which was obtained from PJM.

[135] About 1600 of the 4300 contingencies apply to Classic PJM, but the system operator does not always enforce all of them, see PJM's information on Contingencies at http://www.pjm.com/markets/energy-market/lmp-contingencies.html.

[136] For the analyses reported in this paper, we used the Annual FTR load flow case that PJM posted in February 2007 and the monthly FTR load flow case posted in July 2006.

Daytime electricity demand, however, is typically higher with peak electricity demand reaching about 60 GW in Classic PJM in 2005. It is important to simulate potential NO_X reductions in peak demand hours because the higher demand requires more complete utilization of generating units. If the generating units with low NO_X emission rates were fully utilized to meet demand there would be little flexibility to reduce emissions. In addition, if demand were higher in areas with little generation, or with only costly generation, then higher demand would increase the likelihood of congested transmission lines. If the low-NO_X generation were also located far from high-demand areas then network constraints could similarly limit NO_X reduction potential. In order to simulate high demand conditions, the PJM FTR model was scaled to approximate the higher demand hours studied with the zonal model and the North American Electric Reliability Corporation (NERC) Multiregional Modeling Working Group (MMWG) base-case power flow models were obtained through a Freedom of Information Act filing with the Federal Energy Regulatory Commission.

Three scaled cases were developed that had similar levels of total demand, fossil generation, and NO_X emissions as those observed in historical peak demand hours in Classic PJM. The first of these cases mimicked the historical LMP patterns observed on August 4th at 2pm ("Matched LMPs"). The Matched LMPs case started with two binding constraints in the security constrained optimal power flow. In the second case, the nodal load data were altered until there were 9 initially binding constraints, four of which PJM reported as active on August 4th at 2 pm ("Constraints"). In the third case, there were six initially binding constraints and one of these was observed on August 4th at 2pm. In addition, the Classic PJM fossil units generated 37 GW and emitted 39 tons of NO_X in this base case ("High Fossil Gen"). In the Matched LMP and Constraints cases the initial generation was 34 GW and initial NO_X was 38 tons. On August 4th at 2pm these fossil units generated 35 GW and produced 38 tons of NO_X. The High Fossil Gen case better replicated the observed initial NO_X emissions than the other two cases although the initial generation was higher. It also provided a conservative estimate of potential NO_X reductions because by requiring more generation from the fossil units – there was less under-utilized generation available for redispatch. In addition, the fossil units in Classic PJM also initially generated 37 GW in the highest demand MMWG case (the "Summer" case, discussed below) and the units emitted 43 tons of NO_X. Comparisons between the MMWG Summer case and FTR cases use the High Fossil Gen case for this reason.

The MMWG cases were simpler to use because they did not require scaling. Four cases were used. In the summer case the total electricity demand in Classic PJM was about 59 GW. In the fall and spring cases Classic PJM demand was about 41 GW and 40 GW respectively and in the "low load" case it was about 24 GW. The MMWG cases also contained more information on the capacities of generating units as well as information on how the generators corresponded to the EIA's database of generating units, which still required matching to the EPA's database. The MMWG cases did not include information on contingencies and used different bus numbers and a slightly different network aggregation (or network topology) compared to the FTR cases. The MMWG cases contain information on the entire Eastern Interconnection including the power systems of New York and New England. PowerWorld was used to build an "equivalent" network that contained only the Classic PJM and "electrically equivalent" but simpler approximations of the surrounding systems (see Overbye *et al.* 2004 for another example of using an "equivalenced" system). The imports and exports to and from Classic PJM from the approximated adjoining systems were held constant in the simulations.

For the PowerWorld simulations the DC approximation to the AC load flow was utilized. Both the AC and DC methods solve for the power flows over the network, but the former does not consider reactive power flows or line losses.[137] The literature suggests that DC SCOPF is sufficient for most economic analyses of electricity networks. Schweppe *et al.* (1988) proposed the DC load flow as a tool for economic analysis. Overbye *et al.* (2004) analyzed the accuracy-tractability trade off between using the full AC load flow and the DC SCOPF for LMP studies for the 13,000-bus model of the Midwest U.S. transmission grid. They found that DC SCOPF performed reasonably well: although the power flows were not identical, the DC method identified very similar patterns of constraints and the average LMP only differed by about $2.40/MWh (lower in the DC case). The DC approximation found that some lines were only about 99% loaded while the AC load flow found them to be congested, causing the observed difference in LMPs. Given this finding, any inaccuracies resulting from the use of a DC approximation are likely overshadowed by the use of linear cost curves, the choice only to model Classic PJM and not the entirety of the PJM network, matching the generators to the FTR case buses, and the necessity of scaling the FTR cases to represent peak demand conditions.

5.2. Results and Discussion

Three characteristics of a power system create the flexibility to reduce NO_X emissions (or emissions in generally) through redispatch (herein "NO_X flexibility"). First, for redispatch to be possible at all requires the existence of under- or unutilized generating capacity. Second, NO_X reductions may be possible if some of the underutilized capacity burns natural gas because natural gas units tend to have lower NO_X emission rates than coal and oil units. Third, if the NO_X rates of generators within the same fuel category differ and the low NO_X generation is underutilized then the redispatch of these units could reduce NO_X emissions. The characteristics of capacity, generation, and NO_X emissions Classic PJM suggest that flexibility to reduce NO_X through

---

[137] According to Overbye *et al.* 2004, the major simplifications of the DC power flow are that it 1) ignores the reactive power balance equations, 2) assumes identical voltage magnitudes of one per unit, 3) ignores line losses, and 4) ignores tap dependence in the transformer reactances.

redispatch may be available and the simulations were designed to test whether network constraints limit this potential. The following subsections discuss the results: first, estimates of the maximum technical potential for NO_X reductions by redispatch in PJM and second, estimates of the magnitude of the NO_X prices needed to achieve various levels of NO_X reduction up to that maximum. Because of the temporal- and locational-variations in the impact of NO_X emissions on ozone formation, the results are presented in terms of their temporal and locational characteristics. Comparisons of zonal and PowerWorld simulations are also discussed.

5.2.1. Relevant Background Characteristics of PJM

Both demand and fossil fuel-fired generation in PJM and in Classic PJM were highest during the ozone season (May through September). Table 5-1 displays the average and maximum hourly demand in PJM in 2005 during the ozone season and during the non-ozone season months. The table also shows the average and maximum hourly generation from the fossil-fired generating units used in the simulations ("371 units in Classic PJM").[138] The maximum-demand hour for all of PJM in 2005 occurred on August 3rd at 5 pm. The demand of about 116 GW in that hour, not including the Duquesne Light Company (DUQ) Control Zone, was about 1.6 times that of the average demand in PJM during the ozone season of 2005. The maximum-demand hour for Classic PJM occurred on July 27th at 4 pm with demand of also about 1.6 times that of the average demand in Classic PJM in the ozone season of 2005.

The average hourly NO_X emissions from the units in Classic PJM in 2005 were about 20 tons per hour (Table 5-1). The maximum hourly NO_X emissions in 2005 did not occur during the ozone season in 2005, but occurred in January when the cap-and-trade program for NO_X was not in effect.

---

[138] Our simulations do not model the further possibilities of exchanging hydro or nuclear power for fossil generation – although for nuclear we would expect the possibilities to be small as most nuclear plants are typically run near their full capacity in most hours.

Table 5-1 Average and Maximum demand in PJM and Classic PJM and Fossil Fuel-Fired Generation and Emissions in Classic PJM.

Hourly Data, 2005		**Ozone-Season**	**Off-Season**	**Annual**	
PJM Demand^	avg	74	68	71	*(GW)*
	max	116	97	116	
Classic PJM Demand	avg	36	32	33	*(GW)*
	max	59	46	59	
Classic PJM Fossil	avg	19	16	18	*(GW)*
	max*	36	26	35	
Classic PJM NOx Emissions	avg	19.6	30.0	25.7	*(Tons)*
	max*	44.7	46.2	46.2	

^Does not include the DUQ control area that joined PJM May 1, 2005
*Max from the highest demand hour in Classic PJM in 2005 in the ozone season (7/27/05 16:00) and non-ozone season (1/18/05 19:00) respectively

While total generation in the summer peak hour in Classic PJM was about 28% higher (13 GW) than at the winter peak, the summer peak NO_X emissions were slightly lower, 45 tons in contrast to 46 tons during the winter peak. The increased use of natural gas-fired generation to meet the higher levels of summertime demand can partially explain this: on average, natural gas-fired generators filled about 16% of hourly summer demand but only 10% of hourly demand in the winter. In addition, the average emission rate of coal-fired generation was about 2.15 lbs/MWh in the ozone season and about 4 lbs/MWh outside the ozone season in 2005. The ozone season NO_X price likely explains this lower ozone season emission rate for coal-fired units because, in the absence of a price on NO_X, the NO_X emission rates of coal-fired units would be higher in the summer because emission rates increase with utilization and because of the increased use of less efficient units to fill the higher peak demand.

An important feature of Classic PJM (and all electricity systems in the U.S.) is that even during the hours of the highest peaks in demand, there is generating capacity that is in some form of reserve status and not actually generating electricity. This is the first reason to expect that NO_X reductions through redispatch might be possible. Table 5-2 shows the capacity of the 371 fossil fuel-fired generating units that the simulations redispatched. The total capacity of these units was about 46 GW (or 42 GW if de-rated by the annual forced outage rate for PJM in 2005).[139] The maximum hourly generation from these units during 2005 was about 36 GW, leaving about 6 to 10 GW of capacity that was not generating electricity in the peak hour. Some of this remaining capacity was providing spinning, non-spinning, and supplemental reserve margins for reliability purpose. The simulations assume that units with higher NO_X emission rates that were generating electricity during the peak hours could be exchanged for lower NO_X units in these reserves, at least for short periods of time.

Table 5-2 Capacity and Generation by Fuel-Type in Classic PJM during the 2005 Ozone Season.

Hourly Data, Ozone Season 2005		**Coal**	**Natural Gas**	**Oil**	**TOTAL**	
Capacity	rated	21	15	10	46	*(GW)*
	unforced^	19	14	9	42	
Generation	avg	15	3.0	1.6	19	*(GW)*
	max*	18	10	8.2	36	
NOx Emissions	avg	15.8	1.2	2.6	19.6	*(Tons)*
	max*	20.2	6.9	17.6	44.7	
NOx Emission Rates	avg	2.15	0.78	3.19	2.02	*(lbs/ MWh)*
	max*	2.24	1.37	4.29	2.46	

Fuel Category Designations from the EPA's Clean Air Markets Database
*Max from the highest demand hour in Classic PJM in 2005 in the ozone season (7/27/05 16:00)
^Derated by the equivalent demand forced outage rate for PJM in 2005 (7.3%) (PJM 2006)

Table 5-2 also shows that a mix of fuels were used to generate electricity in Classic PJM and that natural gas generation had the lowest average NO_X rate, about half the average for coal-fired generation. Moreover, and the second reason to expect NO_X flexibility, natural gas-fired capacity represented the largest portion of the unutilized capacity (for both peak and average hours). This likely occurred because the bid-based, security constrained economic dispatch utilized the highest marginal cost units last and

---

[139] Since the annual forced outage rate may be too restrictive, as noted earlier (infra note 129), the range is presented in Table 2.

natural gas-fired units tend to have the highest marginal costs due to natural gas prices (which were particularly high in 2005). For all fuel-types, the generation dispatched to fill peak demand had a higher NO_X rate than that dispatched to fill average demand. This is as expected since there is no differentiation in NO_X pricing between peak and other summer hours and the units pressed into service during peak hours are typically those of all fuel types with lower efficiency (higher heat rates).

For a high NO_X permit price to cause redispatch that reduces NO_X emissions in a given hour, unutilized capacity that is available to generate must have a lower NO_X rate than the original generation used to fill demand. The graphs in Figure 5-1 show cumulative distributions over NO_X emission rate of the generation used to fill demand and the remaining capacity in Classic PJM on August 4th, 2005 at 2 pm (one of the highest demand hours in PJM during 2005). The median NO_X emission rate for this hour was 2.2 lbs/MWh for all units and for coal-fired units. The graphs show that about 42% and 34% of the remaining, undispatched capacity for fossil fuel-fired and coal-fired units, respectively, had a lower NO_X rate than the median for the units used to fill demand in that hour. This provides the third reason to expect NO_X flexibility.

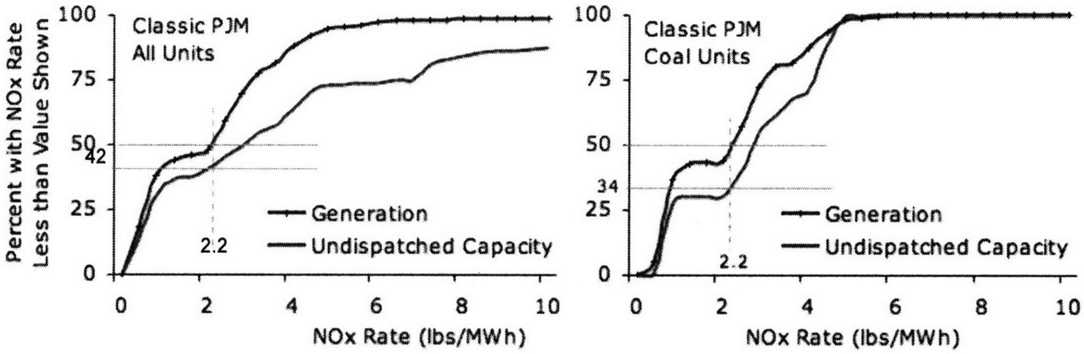

Figure 5-1 Cumulative distributions of generation and undispatched capacity over NO_X Rate in Classic PJM on August 4th, 2005 at 2pm. The graph on the left shows fossil fuel generating units in Classic PJM and that on the right shows only coal units. If a heat rate of 10 mmBTU/MWh is assumed, these NO_X emissions rates translate to the equivalent NO_X emission rate in lbs/mmBTU by dividing by a factor of 10.

This result is fairly consistent across other hours and levels of demand. In hours with lower demand, the median NO_X rate of the units that were generating electricity was slightly higher and that of the undispatched capacity was slightly lower. The availability of relatively low-NO_X capacity, even in high demand hours, suggests that redispatch could reduce NO_X emissions if the economic incentives to do so were in place and network constraints did not prevent the utilization of the lower-NO_X rate generation.

5.2.2. Temporal Variation in Potential NO_X Reductions

The potential NO_X reductions from redispatch vary in time primarily because the total demand for electricity, and the pattern of the demand on the network, varies diurnally and according to the weather.[140] Table 5-3 reports the generation, emissions, and simulated NO_X "reductions" using the zonal model for the 24 hours preceding the peak demand hour of 2005 in PJM, August 4th at 2pm. The range of total hourly generation for the units we considered in Classic PJM was from about 19 GW per hour, which occurred during the middle of the night, to 35 GW on August 4th at 2pm. The range of initial hourly NO_X emissions was between about 19 and 38 tons. The reductions ranged from about 6.1 tons (17%) during the day to 8.4 tons (about 35%) in early morning and late night hours for the transmission constrained estimates (labeled "Transmission").[141] Larger reductions should be possible at night because the network is typically less constrained and less capacity is utilized during the lower demand hours.

---

[140] There will also be some variation due to planned maintenance of facilities, which will be scheduled primarily for other than the peak summer demand season.

[141] Since natural gas prices were high during the summer of 2005, observed emissions, and therefore the simulated reductions, might have been higher than in a more normal year. For comparison, we looked at a peak demand hour of 2001 when natural gas prices were much lower. During this hour, there were about 31 GW of fossil generation in Classic PJM (vice 35 during the peak-hour in 2005) and 51 tons of NO_X emissions (vice 38 tons). The potential unconstrained NO_X reductions were about 16 tons or 32%. Both the initial emissions and NO_X reductions were higher in the 2001 peak-hour than in the near-peak hour in 2005 with the same level of fossil generation (e.g. 8/3/05 20:00); however, the percent reduction was about the same.

Table 5-3 Potential Reductions in NO_X Emissions from Redispatch in Classic PJM using the Zonal Model.

Zonal Model Simulations of Maximum Potential NOx Reductions

	Base Case		***Unconstrained***		***Transmission***		***Unforced Capacity^***		***Only "ON" Units***	
Date	**Generation**	**NOx**	**Reduction**	**%**	**Reduction**	**%**	**Reduction**	**%**	**Reduction**	**%**
8/3/05 14:00	33	35	7.0	20	6.6	19	6.5	18	6.0	17
8/3/05 18:00	33	35	9.2	26	6.1	17	7.4	21	6.1	17
8/3/05 22:00	26	26	10.8	42	6.9	27	9.2	36	6.5	25
8/4/05 2:00	19	19	7.8	42	7.6	41	9.8	52	3.9	21
8/4/05 6:00	23	23	8.6	37	8.4	36	9.3	40	4.5	19
8/4/05 10:00	31	28	7.2	25	6.9	24	6.7	24	4.5	16
8/4/05 14:00	35	38	8.2	21	8.0	21	7.5	20	7.1	19
	(GW)	*(Tons)*	*(Tons)*	*(%)*	*(Tons)*	*(%)*	*(Tons)*	*(%)*	*(Tons)*	*(%)*

^ Capacities were derated by the 2005 demand equivalent forced outage rate for PJM of 7.3% (PJM 2006).

Two additional series of simulations are reported in Table 5-3 in the columns labeled "unforced capacity" and "Only 'ON' units." Both were intended to represent plausible restrictions on the potential to switch generating units that were additional to transmission constraints. In the former, the summertime rated capacities of all generating units were multiplied by a factor of one minus the forced outage rate of PJM in 2005 to represent the possibility that all capacity may not be available at a level of 100% in all hours.[142] The last column represents the case where the low NO_X-emitting units that could substitute for higher NO_X emitting units were limited to those providing spinning reserve services. Of these two further limitations, restricting the pool of exchangeable units to operating units with unused capacity in spinning reserves has the greater effect. Moreover, this effect is significantly greater during non-peak hours than in peak hours. Or, stated differently, most of the NO_X-reducing substitution capability during peak hours comes from units in spinning reserve while most of that during non-peak hours is from units that are not generating at those times.

The available load flow cases restricted the PowerWorld simulations to "generic" hours with different demand, generation, and congestion characteristics (rather than for a series of hours). The results agreed reasonably well with those from the zonal model,

---

[142] PJM (2006), page 244, states that the forced outage rate for PJM in 2005 was 7.3% for all generating units. This rate does vary by type of generating unit (steam units have the highest outage rate and combined cycles the lowest of the fossil-fuel fired units). In this analysis, the capacities of all generating units were scaled by a factor of 0.927.

although the zonal simulations tended to be slightly optimistic compared to the PowerWorld simulations. Table 5-4 shows PowerWorld optimal power flow results for high NO_X prices of $125,000/ton for cases with varying levels of demand and generation from the set of redispatched fossil units. NO_X prices above $100,000/ton caused only small additional reductions in NO_X emissions (see Table 5-7). In Table 5-4 the base case was the result of OPF dispatch with assumed NO_X prices of $2000/ton (indicated by "2k") to roughly represent the observed NO_X prices of between 2000 and 3000 $/ton in the summer of 2005. (The security constraints (SCOPF simulations) did not alter the magnitude of potential reductions but did cause two tons of additional base-case NO_X emissions – see Table 5-6.)

Table 5-4 suggests that the maximum physical reductions depend on the initial levels and patterns of demand and are between about 6 and 8 tons hourly (between about 13% and 30%) in Classic PJM. The MMWG cases yielded the most conservative estimates of the potential NO_X reductions of about 6 tons per hour.

Table 5-4 Potential Reductions in NO_X Emissions from Redispatch in Classic PJM using PowerWorld optimal power flow.

PowerWorld Simulations of Maximum Potential NOx Reductions

		Base Case		*Unconstrained*		*Trans. Const.*	
		Generation	**NOx**	**Reduction**	**%**	**Reduction**	**%**
Peak Demand	**Matched LMP**	34	35	8.2	*23*	8.0	*23*
	Constraints	34	35	7.4	*21*	7.2	*21*
	High Fossil Gen	37	39	7.5	*19*	6.4	*16*
	MMWG Summer	37	43	5.9	*14*	5.8	*13*
Average and Low Demand	**Avg Demand**	19	20	12	*60*	12	*60*
	MMWG Low Demand	14	16	6.9	*43*	6.9	*43*
	MMWG Spring	24	28	7.5	*27*	7.5	*27*
	MMWG Fall	23	26	7.7	*30*	7.7	*30*

Comparable zonal and PowerWorld cases are the peak MMWG Summer case and August 4th at 2pm and the average MMWG Fall case and August 4th at 6am. The generation in the two peak cases was slightly different (37 GW in the MMWG Summer case compared to 35 GW on Augst 4th at 2pm). The potential reductions in the MWMG case were only about 6 tons (from 43 tons) compared to 8 tons in the zonal model simulations (from 38 tons). The NO_X reductions represented a change in average emission rate by 13% (from 2.3 to about 2.0 lbs/MWh) in the MMWG Summer case and by about 18% (from 2.2 to about 1.8 lbs/MWh) in the zonal simulation of August 4th at 2pm. The generation in both the average cases (MMWG Fall case and August 4th at 6am) was 23 GW. The zonal simulation reduced NO_X emissions by 8.4 tons (from 23 tons) compared to 7.7 tons in the MMWG Fall case (from 26 tons initially). The reduction in initial average NO_X emission rate was 35% in the zonal simulation (from 2.0 to 1.3 lbs/MWh) and 30% in the MMWG Fall simulation (from 2.3 to 1.6 lbs/MWh). The zonal simulations were slightly optimistic compared to the PowerWorld simulations (this is discussed further in the Section 5.2.4).

Recent actions taken in the OTC States suggests that the magnitude of potential NO_X reductions from redispatch is nontrivial. A recent OTC Memorandum of Understanding (MOU) signals an intention by the signatory states to reduce emissions on high electricity demand days.[143] Four of the signatory states are in the Classic PJM region and the MOU requires these states to make total daily NO_X reductions of about 72 tons on high electricity demand days, an average of 3 tons per hour over a 24-hour period.[144] Given that 6 tons of reductions are available from redispatch in the highest demand hours, the potential reductions from redispatch are about twice the targets for reducing NO_X emissions.

---

[143] The states agreed to make the reductions beginning in 2009 and no later than 2012. See, OTC's "Memorandum of Understanding Among the States of the Ozone Transport Commission Concerning the Incorporation of High Electrical Demand Day Emission Reduction Strategies into Ozone Attainment State Implementation Planning," March 2, 2007. The MOU does not fully define a high electricity demand day, but some related analysis suggests that these are the days on which the high demand requires peaking units that typically generate in less that 10% of annual hours to generate power (NESCAUM 2006)

[144] The four signatory states that are in the Classic PJM area are DE, MD, NJ, and PA. The other signatory states are CT and NY.

A major difference between the FTR and MMWG cases was the pattern of the loads on the network. The locations of the loads on the network partially determine which generators are dispatched to fill them because of transmission constraints. Even if the overall magnitude of the demand is the same in two cases, different generators might be dispatched to fill them and this causes differences in NO_X emissions. For example, the same set of fossil units generated 37 GW in the "High Fossil Gen" FTR case and in the MMWG Summer case, but the NO_X emissions in the "High Fossil Gen" base case were 39 tons compared to 43 tons in the MMWG Summer case. Because the units have the same NO_X emission rates in both cases, this demonstrates that different units were initially dispatched to fill demand in each case due to differences in the patterns of demand on the network. The resulting redispatch and potential NO_X emissions were also different as a result. The scaling process exaggerated provided, at best, a rough approximation of nodal peak load patterns in the FTR cases so the MMWG summer case is likely more representative of peak demand conditions in Classic PJM. Notably, more lines were congested in the FTR simulations with the scaled loads (6 lines in the "High Fossil Gen" case) compared to the MMWG Summer case (only 2 lines) with the same total demand characteristics. This suggests that "congestion" *per se* does not limit the flexibility to reduce emissions through redispatch and that the nodal pattern of demand relative to the locations of the generators has a greater effect.

5.2.3. The impact of network constraints on potential NO_X reductions

The most striking feature of the results is that transmission constraints do not significantly reduce potential NO_X emissions reductions from redispatch in Classic PJM. There are three primary reasons. The first is related to the spatial heterogeneity in the low and high NO_X generating units in PJM. High NO_X units are not mostly in one area of PJM and low NO_X units in another; they tend to be located together *within* the zones created by transmission constraints. This is particularly important in high demand hours.

In these hours congestion is less of a problem if local demand is predominantly filled by local generation. If there is significant local NO_X-rate heterogeneity then NO_X emissions can be reduced without substantial increases in the utilization of transmission lines.

Figure 5-2 suggests that there is local heterogeneity in the NO_X emission rates of generating units. The figure shows distributions of generation over NO_X rate for all units, for only coal units, and for all units located in Middlesex County, NJ. The two lines represent generation as observed and as simulated when all units have been redispatched to minimize NO_X constraints using the zonal model to simulate August 4th, 2005 at 2 pm. As would be expected, the range of the distribution of generation across NO_X emission rates is similar among the three panels and it is not drastically altered in the simulations.

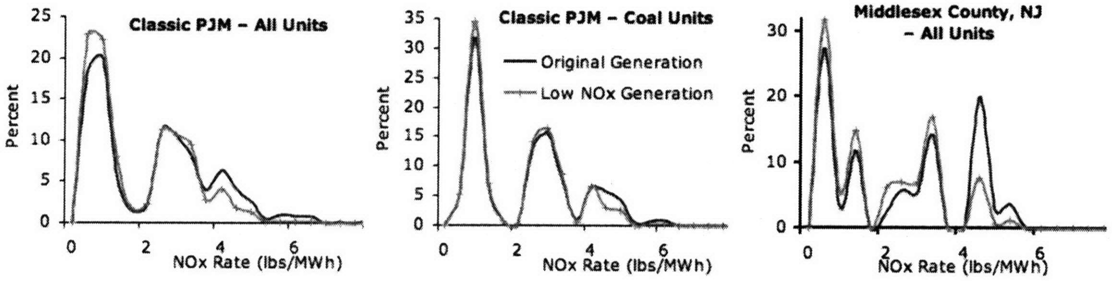

Figure 5-2 Distributions of generation over NO_X rate for all units in Classic PJM, for coal units in Classic PJM, and for all units in Middlesex County, NJ. If a heat rate of 10 mmBTU/MWh is assumed, these NO_X emissions rates translate to the equivalent NO_X emission rate in lbs/mmBTU by dividing by a factor of 10. Redispatched cases are from the zonal model for August 4th, 2005 at 2 pm.

The trimodal distribution that is observed for all units in Classic PJM is as true for coal units as it is for the entirety of units and it is still evident in the distribution for Middlesex County. The main effect of the Low NO_X Case is to shift generation from the high (> 4 lbs/MWh) part of the distribution to the two lower modes for all three cases. The shift is particularly evident in Middlesex County where the share of generation in the high emission rate segment is reduced by about two-thirds. In cases like this, which occur in many of sub-regions of Classic PJM at the county-scale, transmission constraints are simply not a problem.

The second reason for the small effect of transmission constraints is that, to the extent low-NO_X generation is located at one end of a congested line, it tends to be on the high-LMP side of the constraint. For example, the capacity-weighted average NO_X emission rate of the units on the low-LMP side of the frequently constrained 10THST to OST line was 3.1 lbs/MWh in the summer of 2005, while that on the high-LMP side was 1.8 lbs/MWh. On August 4th, 2005 at 2 pm, the generation on the low-LMP side of this constraint had an average NO_X rate of 2.6 lbs/MWh and that on the high-LMP side an average NO_X rate of 1.7 lbs/MWh. Anything that would increase the use of unused low-NO_X generation on the high-LMP side of the constraint in place of the higher-NO_X generation on the low-LMP side will relieve the transmission constraint. Here again, the transmission constraint was not a problem because the NO_X-reducing exchange creates a flow in the opposite direction.

The third and final reason is that NO_X reducing substitutions involve small amounts of generation, especially in the peak hour. In peak demand hours in PowerWorld and the zonal model, the simulations exchanged about 4.5 GW of generation to reduce emissions to the physical limit, within a set of units contributing about 35 GW total. In the average demand hour, the simulations exchanged about 8.5 GW of generation of about 20 GW total.

5.2.4. Locational Variation in NO_X Reductions

The location, in addition to the time, of NO_X reductions affects their impact on ozone formation. One of the first criticisms of the cap-and-trade approach was that "hotspots" could result because these programs have not traditionally captured time and locational variations of the impacts of emissions on air quality standards. These hotspots, which have not been shown to occur in any of the currently implemented cap-and-trade programs, would occur when sources in an environmentally sensitive area chose to buy permits for their pollution, rather than taking actions that resulted in abatement.[145] This motivates the question of whether the redispatch of units to reduce NO_X is accompanied by substantial increases in NO_X emissions in some geographic areas.

It is certainly true that on the level of individual plants, some locations will produce more and some will produce less NO_X as a consequence of redispatch. But at a higher level of aggregation it is not necessarily true that the redispatch, which results in a net reduction of NO_X, will result in areas with significantly higher NO_X emissions. Table 5-5 shows the observed NO_X emissions by county for August 4th at 2 pm and the base case NO_X in the MMWG Summer case. It also shows the changes in NO_X and generation due to redispatch subject to network transmission constraints in the zonal model and the reductions in the MMWG summer case with NO_X prices of \$100,000/ton. The table shows only those counties in which the redispatch changed NO_X emissions by at least 200 lbs.

In the MMWG summer case, emissions increased the most in Middlesex County, New Jersey as a consequence of the increased output of one generating unit. The same unit reduced its output as a consequence of redispatch in the zonal model of August 4th at 2 pm, so emissions in Middlesex County decreased in the zonal simulation. Emissions increased the most in the August 4th, 2pm zonal simulation in the District of Columbia. Again this increase occurred because of the increased utilization of one generating unit. In the MMWG summer case the redispatch caused the unit to generate less, decreasing emissions in DC. The air quality consequences of these changes would ultimately depend on the meteorology and atmospheric chemistry conditions at the times they occurred.

---

[145] For a summary of analyses of these issues see Swift (2004).

Table 5-5 Original emissions and changes in at the county-level for simulated redispatch subject to network constraints for the MMWG Summer OPF simulation in PowerWorld and on August 4^{th}, 2005 at 2 pm in the zonal model. The chart shows counties that had a net change in NO_X of at least 200 lbs.

		PowerWorld MMWG Summer		Zonal August 4th, 2 pm	
State	**County**	**NOx**	**Change in NOx**	**NOx**	**Change in NOx**
NJ	Burlington	4580	-4234	2553	-1557
PA	Bucks	3012	-2240	335	-257
NJ	Hudson	5624	-1808	5370	-3258
MD	Harford	1841	-1035	1146	-749
MD	Talbot	1017	-1017	0	0
NJ	Essex	1044	-963	719	-337
PA	Philadelphia	1628	-898	546	32
PA	Clearfield	2206	-674	1464	-967
NJ	Cape May	1969	-488	1752	-1134
MD	Prince Georges	6375	-357	5283	-715
PA	Northampton	3452	-208	6304	-1754
DC	DC	1470	0	613	1011
PA	Venango	27	89	81	213
PA	Delaware	3185	126	3141	257
NJ	Gloucester	463	202	427	-3
MD	Baltimore	2050	206	2605	-1451
PA	Union	0	223	0	0
MD	Dorchester	311	558	744	-595
NJ	Middlesex	3034	1782	4651	-1716
			lbs		

The magnitudes of the increases in NO_X were generally small. Emissions increased more than 200 lbs in only 5 counties in the MMWG Summer case and in 3 counties in the August 4^{th}, 2 pm zonal simulation. In both simulations, the redispatch increased emissions in 19 of the 57 total counties.

As Table 5-5 suggests, the initial emissions and the changes from redispatch were different in the PowerWorld and zonal models. The nodal load data for historical hours, such as August 4^{th} at 2pm, are not available. This makes simulating historical hours with PowerWorld difficult and makes it a challenge to determine whether the differences in the results of the zonal model and PowerWorld simulation were more a result of the simulation method or of initial conditions. The significant differences between the PowerWorld simulations that used the FTR and scaled FTR cases and the MMWG cases suggest that initial conditions partially determine the potential NO_X reductions from redispatch.

A rough comparison of the changes in generation from units in the PowerWorld MMWG Summer case and the zonal August 4th, 2pm case does suggest, however, that the zonal simulations allowed many more substitutions than did the PowerWorld simulations. For example, there were 191 units that generated in the base cases for both of these simulations. Of these 191 units, 95 units had similar levels of initial generation in both cases (less than 11.5 percent different, the median difference in generation between the two cases was 11.5 percent).

Although the extent to which these 95 units can be redispatched depends on the initial states of other generating units and on the pattern and magnitude of nodal loads, it is somewhat telling to compare the changes in generation of these units between the two simulation methods. In the PowerWorld MMWG Summer base case, the 95 units generated about 19.1 GW and in the zonal base case they generated about 19.2 GW. After redispatch, the units generated 19.4 GW in the MMWG Summer case and 19.1 GW in the zonal case. The redispatch changed only 3 units' output by more than 20% in the MMWG Summer simulation, while it changed 23 units' output by more than 20% in the zonal simulation. This suggests that the constraints limiting the exchange of generation from units in the PowerWorld simulations are more stringent than those in the zonal simulations and that the PowerWorld simulations give a more conservative estimate of the potential NO_X reductions from redispatch.

5.2.5. SCOPF Simulations and Emissions-security Tradeoffs

The results discussed this far suggest that the redispatch of power plants in Classic PJM can cause nontrivial reductions in NO_X emissions from dispatch of about 15% in peak demand hours and 30% in average demand hours. Network congestion does not drastically limit the potential reductions, but the magnitude and pattern of initial nodal load and generation partially determine the potential NO_X reductions and so does a realistic representation of the network. Security constraints are an additional restriction on the operation of the electricity grid.

Data on security constraints were available for the FTR base case power flows. In Classic PJM, the data included 1455 contingency constraints. It is possible that the consideration of this entire set of contingency constraints is overly restrictive. PJM reports on its website that they do not always enforce all contingency constraints and their operating procedures allow for the system operators to use their judgment with regard to whether lines can be overloaded.[146] The addition of security constraints not only changes the potential NO_X reductions from redispatch but also the base case NO_X emissions. Table 5-6 shows the OPF and SCOPF simulations in the FTR High Fossil Gen and Average Demand cases with the entire set of 1455 contingency constraints.

Table 5-6 OPF and SCOPF Simulations for the FTR High Fossil Gen and Average Demand cases.

	NOx		OPF			SCOPF	
	Price	NOx	Reduction	%	NOx	Reduction	%
High Fossil Gen	2k	39	--	--	41	--	--
	100k	33	6.5	*17*	34	6.9	*17*
Average Demand	2k	20	--	--	17	--	--
	100k	8	11.3	*58*	12	5.7	*33*
	$/ton	*Tons*		*%*	*Tons*		*%*

In the High Fossil Gen case the *initial* NO_X emissions (in the $2000/ton NO_X price base case) *increased* from 39 to 41 tons with the addition of security constraints. The potential NO_X reductions were 17% in both the OPF and SCOPF simulations. In the Average Demand the opposite was true: the *initial* NO_X emissions *decreased* from 20 to 17 tons with security constraints and the potential reductions from redispatch were smaller. Because the impact of security constraints on the initial NO_X emissions and

---

[146] See PJM's information on Contingencies at http://www.pjm.com/markets/energy-market/lmp-contingencies.html.

potential reductions varies with the level of electricity demand, there is not a clear tradeoff between the goals of security and environment in the case of NO_X emissions.

5.2.6. NO_X Prices Needed to Encourage Redispatch

Table 5-7 shows the relationship between NO_X prices and potential reductions in NO_X emissions for the PowerWorld simulations. All simulations economically dispatched the generators in Classic PJM (minimized total operating costs) for ranges of NO_X prices in the average and peak demand hours using the cost curves discussed in Section 5.1.4. The simulations suggest that even in the unconstrained case in the average demand hour, NO_X prices of about $50,000/ton would be necessary to obtain substantial reductions.[147] The NO_X reductions at $50,000/ton in both the average and peak demand cases were similar, about 5 or 6 tons. In the average demand hour, higher NO_X prices caused further reductions by increasing generation from natural gas. In the peak demand case, these natural gas units were already generating; there was less excess capacity to exchange.

Table 5-7 Results of the PowerWorld simulations for a range of assumed NO_X permit prices. Reductions (absolute and percentages) are calculated from the $2000/ton (2k) NO_X price case in the corresponding panel.

	MMWG Summer						***MMWG Fall***					
NOx	**Unconstrained**			**OPF**			**Unconstrained**			**OPF**		
Price	**NOx**	**Reduction**		**NOx**	**Reduction**		**NOx**	**Reduction**		**NOx**	**Reduction**	
2k	43	--	--	43	--	--	26	--	--	26	--	--
10k	40	3.0	*7*	40	3.0	*7*	24	1.8	*7*	24	1.8	*7*
20k	38	4.5	*11*	38	4.5	*10*	23	2.9	*12*	23	3.0	*12*
50k	37	5.4	*13*	38	5.2	*12*	19	6.2	*24*	19	6.1	*24*
100k	37	5.8	*14*	37	5.6	*13*	18	7.4	*29*	18	7.5	*29*
125k	37	5.9	*14*	37	5.8	*14*	18	7.7	*30*	18	7.7	*30*
$/ton	*Tons*	*Tons*	*%*	*Tons*	*Tons*	*%*	*Tons*	*Tons*	*%*	*Tons*	*Tons*	*%*

---

[147] If the NO_X emission rate of the marginal generating unit were 3 lbs/MWh then a $20,000/ton NO_X price would add (roughly) $30/MWh to the locational price for electricity. If the marginal generating unit had a NO_X rate of only 0.5 lbs/MWh, the NO_X price would only add about $5/MWh to the locational price for electricity.

The average costs of NO_X abatement from redispatch were calculated by dividing the total change in fuel and O&M costs by the NO_X reductions between the different NO_X price simulations and the $2000/ton base case simulation. Table 5-8 shows the average abatement costs for the MMWG Summer and Fall OPF simulations. The average costs of abatement at NO_X prices below about $20,000/ton are roughly half of the NO_X prices.[148]

Table 5-8 NO_X emissions, abatement, and average abatement costs for NO_X prices between $2000/ton and $100,000/ton in the MMWG Summer and Fall simulations.

MMWG Summer Case **NOx Price**	**NOx**	**Abatement**	**Percent Abatement**	**Average Cost of Abatement**
2k	42.7	--	--	--
10k	39.8	3.0	7	5741
15k	39.2	3.6	8	6635
20k	38.3	4.5	10	9203
25k	38.1	4.6	11	9397
30k	38.0	4.7	11	9778
50k	37.5	5.2	12	15061
100k	37.2	5.6	13	21414
$/Ton		*Tons*	*%*	*$/Ton*

MMWG Fall Case **NOx Price**	**NOx**	**Abatement**	**Percent Abatement**	**Average Cost of Abatement**
2k	25.6	--	--	--
10k	23.8	1.8	7	5792
15k	23.1	2.5	10	7524
20k	22.6	3.0	12	10844
25k	22.0	3.6	14	14911
30k	21.3	4.3	17	18964
50k	19.5	6.1	24	32913
100k	18.1	7.5	29	44218
$/Ton		*Tons*	*%*	*$/Ton*

Figure 5-3 shows the marginal and average abatement cost curves for the same simulations. Especially in the MMWG Summer case, the costs increase steeply as the simulations approach the maximum potential NO_X reductions. The costs are also similar for the MMWG Summer and Fall cases for abatement up to about 10%, at which point the costs increase more quickly in the higher demand MMWG case.

---

[148] Personal conversations with industry representatives suggests that this was they expected.

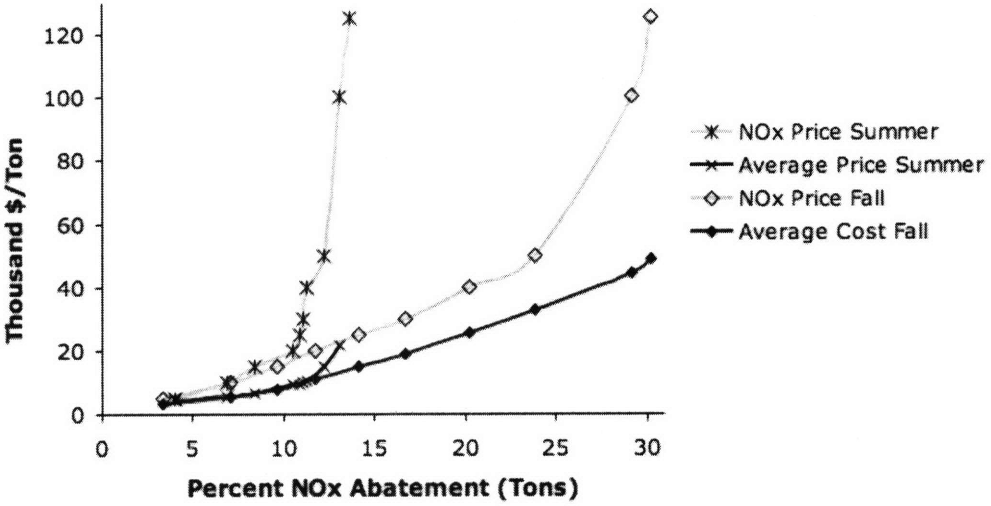

Figure 5-3 Marginal and average abatement cost curves for the MMWG Summer and Fall OPF simulations.

The OTC MOU again provides a point of comparison (see Section 5.2.2). The MOU does not require specific actions to reduce the peak demand day NO_X emissions and it notes that the reductions could come from controls on peaking units or through other measures like energy efficiency or demand response. As an example of action that states could take to control emissions from power plants on peak electricity demand days, the EPA calculated that the average abatement costs of installing water injection NO_X control technology on peaking units in the Northeastern U.S. would be about $158,000/ton to reduce NO_X by about 0.23 tons per day over a 12-day, high-electricity-demand period for each unit that installed the technology.[149] The same EPA analysis estimated average costs of installing SNCRs on uncontrolled coal plants of $18,000/ton to reduce NO_X by about 2.2 tons per day per unit over the same 12-day period. Redispatch appears preferable on a cost per ton basis to controlling NO_X emissions from infrequently used peaking units, although other control options may also be available.

---

[149] EPA Clean Air Markets Division presentation by Chitra Kumar, "High Electricity Demand Day Attainment Strategies for the OTC," December 6, 2006.

One of the benefits of time varying NO_X prices is that the control decisions could be made through decentralized market incentives rather than by regulatory fiat. Another related benefit is that, with the incorporation of air quality forecasting, these costly reductions could come during the times and locations that would most likely impact ozone formation in critical areas – rather than from a specific, predefined set of generating units. For comparison to these cost examples, Mauzerall *et. al.* (2005) estimated the damages of ozone per incremental ton of additional NO_X emissions to be between about \$13,000 and \$64,000 per ton.[150]

5.2.7. NO_X Reductions from Changes in Fuel Use

A NO_X price of about \$50,000/ton was required to reverse the merit order of typical coal and gas generating units given summer of 2005 fuel prices. The exact NO_X price for particular plants depended on heat rates, NO_X rates, and fuel prices. For example, a coal and a natural gas generating unit with only fuel and NO_X costs would generate one megawatt of electricity for the same cost if NO_X prices were \$52,000/ton, coal prices were \$2.5/mmBTU, natural gas prices were \$9/mmBTU and the units had respective NO_X emission rates of 0.4 and 0.15 lbs/mmBTU and heat rates of 10 mmBTU/MWh. The NO_X price to cause substitution would decrease with increases in the coal unit's NO_X rate or the natural gas unit's heat rate.

The heterogeneity in the emission rates and efficiency of power plants (Chapter 4) caused substitution between coal, natural gas, and oil to occur for prices that ranged between \$2000 and about \$100,000/ton. NO_X prices of \$20,000/ton did not cause any change in coal generation in the MMWG Summer simulation, but natural gas generation did replace oil generation (Table 5-9). Some substitution of natural gas generation for coal occurred in the MMWG Summer case at NO_X prices of \$50,000/ton and more at

---

[150] Mauzerall *et. al.* (2005) page 2863. Estimates converted from 1995 to 2005 dollars with a Consumer Price Index conversion factor of 0.78.

higher prices. A small amount of coal to natural gas substitution occurred at lower prices in the lower demand MMWG Fall simulation and much more substitution occurred at the higher prices.

Table 5-9 Changes in generation from coal, oil, natural gas and municipal solid waste (MSW) caused by NO_X prices of $20,000/ton (20k), $50,000/ton (50k), and $100,000/ton (100k) compared to the base case generation ($2000/ton NO_X prices).

	Base Case	Change in Generation		
MMWG Summer	**Generation**	**20k**	**50k**	**100k**
Coal	17996	0	-196	-425
Oil	7185	-182	-128	-104
Natural Gas	10377	198	409	670
MSW	225	0	-60	-115
MMWG Fall				
Coal	16461	-132	-1900	-2965
Oil	2217	-142	393	844
Natural Gas	4130	284	1548	2233
MSW	143	-10	-40	-112
		MW		

The substitution between fuels does not explain all of the simulated NO_X reductions. Within-fuel substitutions also caused NO_X reductions. For example in the $100,000/ton MMWG Fall simulation, coal generating units reduced their output by a total of 3487 MW but the net change in coal generation was a decrease of 2965 MW (Table 5-9). This meant that some coal units increased their output, by a total of about 522 MW. Likewise, there was also substitution within natural gas generating units of about 1000 MW in the same simulation.

5.2.8. Distributional Effects

The distributional effects of a high, short-term NO_X price (that a differentiated cap-and-trade program could cause) for companies that own generating units depend on the electricity demand in the hours in which the higher NO_X price applies and on the initial allocation of NO_X emission permits. The initial allocation of permits to emission sources determines the distributional consequences of any cap-and-trade program.[151] If transaction costs are low, the initial allocation of permits does not affect a cap-and-trade program's ability to achieve the required emission reductions at least cost, but the initial allocation does determine which sources will be required to purchase permits to cover their emissions and which will have extra to sell.[152]

The distributional effects of a high NO_X price for power plants in the redispatch simulations depend on the level of electricity demand because both electricity demand and NO_X prices affect the marginal generating units – the units that set the wholesale electricity price at different network nodes.[153] When a generating unit is at or near the margin it does not earn large profits. In general, coal power plants have the lowest marginal costs and the largest profit margins of all the fossil fuel generators, which they use to cover fixed costs. But a high NO_X price could cause coal to become the marginal fuel, decreasing the profit margins of coal plants and increasing those of the natural gas plants with lower NO_X emissions. For any given NO_X price, the level and pattern of electricity demand will determine the marginal generating units. If the marginal generator burns fossil fuel, which is typically the case, the price of electricity will increase to reflect the costs associated with NO_X emissions. Companies that own nuclear and hydro plants will enjoy larger profits as the margin between the price of electricity and their operating costs grows.

A rough measure of the effects of redispatch on the changes in relative profits of companies that own the redispatched fossil units is the change in share of the total profits

---

[151] For example analyses of the effects of distributional consequences on the potential and actual implementation of cap-and-trade programs see Atkinson and Tietenberg (1982) who analyzed the distributional consequences and other attributes of cap-and-trade program design for particulate control in St. Louis using simulations. Joskow and Schmalensee (1998) used empirical methods to study the political economy of the initial allocation of permits in the U.S. Acid Rain Program. See generally Tietenberg (2006).

[152] See Stavins (2005) for a discussion of the impact of transaction costs on the efficiency implications of the initial allocation of permits in a cap-and-trade program.

[153] The number of marginal generating units in a power network depends on whether or not there is congestion. If no lines are congested there will only be one marginal generating unit.

between two NO_X price scenarios. The simulations of Classic PJM allow only a rough characterization of the distributional effects of a differentiated cap-and-trade program. The simulated LMPs only roughly reproduced the magnitudes and patterns of observed LMP (Appendix E) and the cost curves were also simplified. This meant that the simulated profit margins were not necessarily accurate in an absolute sense. Also, whether or not the differentiated program would require some companies to buy additional permits would depend on other control decisions throughout the ozone season, on which days and how many days the high exchange rate applied, and on whether the exchange ratio could be relaxed on days when ozone pollution was not a problem.

Table 5-10 shows the profit shares of the ten companies with the largest shares in the MMWG Fall base case simulations. It also shows two scenarios of the potential changes in profit shares of these companies for the $50,000/ton NO_X-price redispatch. The first scenario assumed that all companies purchased (from an imaginary bank) enough permits at the prevailing price to cover all emissions in that hour. The second scenario assumed that each generating unit was given an allocation of 76% of its base case emissions because the total NO_X reductions in the MMWG Fall $50,000/ton NO_X-price simulation were 6.1 tons (24%). In this case, the owners would only have had to purchase permits if their emissions exceeded 76% of the units' base case emissions and they could have sold any additional permits at the prevailing price of $50,000/ton. The results of the same "allocation" scenarios are shown for the MMWG Summer case for the same companies (those with the top ten profit shares in the MMWG Fall case). The allocation for the MMWG summer case was 88% of base case emissions because the $50,000/ton NO_X price caused emission reductions of 12%.

Even these rough scenarios demonstrate that the level of demand and the allocation of permits would determine the relative effects of the differentiated cap-and-trade program on the companies. For example, the profit shares of the ten companies were different in the base case simulations of the MMWG Fall and Summer cases. This demonstrates the effects of the level of demand and is expected because of the varied ownership of different types of generators and because the high demand causes higher electricity prices and increased generation from natural gas and other peaking units. Although their profit shares decreased between the Fall and Summer cases, the top two companies were the same in both the base case simulations. Eight of the top ten companies remained in the top ten (see the "ranking" columns in Table 5-10).

Table 5-10 Profit shares and changes in profit shares due to \$50,000/ton NO_X prices assuming the owners of generating units must purchase permits to cover all emissions at the prevailing price ("Profit Share NO_X at \$50,000/ton") and assuming they received enough permits to cover 76% of their base case emissions in the Fall case and 88% of their base case emissions in the Summer case ("Profit Share … with Allocation").

	MMWG Fall Case (OPF Simulations)						***MMWG Summer Case (OPF Simulations)***					
Owner	Rank in MMWG Fall Base Case	Profit Share Base Case (NOx at $2000/ton)	Profit Share NOx at $50,000/ton	% *Change*	Profit Share NOx at $50,000/ton with Allocation	% *Change*	Rank in MMWG Summer Base Case	Profit Share Base Case (NOx at $2000/ton)	Profit Share NOx at $50,000/ton	% *Change*	Profit Share NOx at $50,000/ton with Allocation	% *Change*
C1	1	17.7	23.3	*32*	18.3	*3*	1	13.7	13.3	*-3*	12.7	*-8*
C2	2	15.9	9.7	*-39*	11.5	*-27*	2	8.3	3.9	*-52*	4.9	*-40*
C3	3	10.7	11.4	*7*	8.4	*-21*	6	4.6	3.5	*-24*	2.9	*-36*
C4	4	8.5	9.1	*6*	6.7	*-22*	9	4.2	3.1	*-26*	2.6	*-37*
C5	5	6.4	6.7	*5*	4.9	*-23*	11	3.0	2.7	*-11*	2.3	*-25*
C6	6	5.9	3.6	*-39*	5.6	*-6*	7	4.5	3.4	*-25*	4.2	*-7*
C7	7	5.6	5.9	*7*	4.4	*-21*	5	5.2	5.2	*0*	4.4	*-16*
C8	8	5.4	4.2	*-22*	6.6	*21*	8	4.3	6.4	*49*	8.0	*85*
C9	9	3.7	5.5	*49*	5.5	*50*	4	5.8	7.6	*32*	7.8	*35*
C10	10	2.8	3.0	*4*	2.2	*-24*	16	2.1	2.1	*2*	1.8	*-14*
		(% of Total Profits)		%	*(% of Total Profits)*	%		*(% of Total Profits)*		%	*(% of Total Profits)*	%

The profit shares of some companies decreased with the higher NO_X price in both the Fall and Summer simulations and regardless of the allocation of permits (e.g. C2 and C6) and increased in all scenarios for one company (e.g. C9). The results for other companies were ambiguous. For example, regardless of the allocation scenario, C1's profit share increased in the Fall case with the higher NO_X price but decreased in the Summer case. The profit share of C8 increased dramatically in the MMWG Summer case with the high NO_X price and in the MMWG Fall case with the initial allocation, but decreased in the Fall case without allocation.

Another way to analyze the distributional effects of a differentiated cap-and-trade program is to compare the differentiated approach to a more stringent aggregate cap. Three scenarios were used to make this comparison. In all scenarios, a day was constructed out of 12 peak hours (using the MMWG Summer simulations) and 12 average hours (using the MMWG Fall simulations). Then in the first scenario ("Stringent Aggregate Cap") a NO_X price of $5,000/ton was applied in all hours to approximate the effects of a more stringent aggregate cap. The second scenario ("Differentiated Peak") approximated the differentiated program by applying a NO_X price of $50,000/ton to four of the peak demand hours, while using the base-case price of $2,000/ton in the other 8 peak demand hours and in the 12 average demand hours. The third scenario ("Differentiated Average") approximated the differentiated program by applying a NO_X price of $50,000/ton to four of the average demand hours, while using the base-case price of $2,000/ton in the other hours. The two differentiated scenarios represent the possibility that NO_X reductions could be needed to reduce ozone at different times of the day.

All three scenarios caused similar levels of daily NO_X reductions. In the base case, the generators emitted about 820 tons of NO_X during the day. The Stringent Aggregate Cap and Differentiated Peak scenarios reduced the NO_X emissions by about 21 tons per day and the Differentiated Average scenario reduced NO_X by about 24 tons.

In all three scenarios, each companies' profits from its fossil generators[154] were compared to the base case of the $2000/ton NO_X price applied to all hours (12 peak, 12 average) assuming that the companies were required to purchase the permits necessary to cover their emissions at the prevailing price in the hour. The Stringent Aggregate Cap and the Differentiated Peak scenarios both increased the total aggregate profits from the fossil units of these companies by about $5.4 million (12%) and by about $6.6 million

---

[154] The profits considered were total revenues from electricity generation less fuel costs, variable O&M costs, and NO_X costs (assuming permits for all NO_X emissions were purchased at the assumed price).

(14%) respectively. The simulated cap-and-trade programs increased the aggregate profits because the increases in locational marginal prices more than compensated for the increases in costs caused by the NO_X prices.

In the Differentiated Average Scenario, the overall profits decreased by about $1.2 million (3%). The decrease in this scenario occurred because the $50,000/ton NO_X price in the average demand hour caused much more substitution from coal to natural gas than the same price applied in the peak demand hour. This substitution caused coal to become the marginal fuel and therefore caused a larger portion of the generation to have a smaller profit margin (the difference in costs between natural gas and coal generators is larger than the difference between coal generators with different NO_X rates). This suggests that although the application of higher exchange rates for permits during average hours could cause more NO_X reductions than the same exchange rate in peak demand hours, it would also be more costly for the owners of coal plants if the higher exchange rates for permits applied during average hours. Ultimately, however this result would depend on the initial allocation of permits.

All three scenarios also created winners and losers. The Stringent Aggregate Cap scenario decreased the profits of 13 companies that owned fossil generation, compared to the base case, and increased profits of 51 (of 65 total). The Differentiated Peak scenario decreased profits for 11 companies and increased profits for 53. The Differentiated Average scenario decreased profits for 29 and increased profits for 26 companies. Table 5-11 shows the base case profits and changes in profits for the three companies in each scenario that gained and lost the most (there was some overlap so only ten companies are shown in the table).

Table 5-11 Changes in profits from fossil generators for three cap-and-trade program scenarios for all companies in Classic PJM with fossil generation as well as for the companies most affected.

		Stringent Aggregate Cap Scenario			Differentiated Peak Scenario			Differentiated Average Scenario		
	Base Case Profit	**Change in Profit**		*Percent*	**Change in Profit**		*Percent*	**Change in Profit**		*Percent*
Total	46	5.4		*12*	6.6		*14*	-1.2		*-3*
Average	0.7	0.08		*12*	0.10		*14*	-0.02		*-3*
Companies										
C114	1.9	1.54	g	*79*	0.61		*31*	-0.10		*-5*
C115	1.8	1.36		*77*	0.27		*15*	-0.33		*-19*
C8	4.9	0.35		*7.1*	1.36		*28*	-0.08		*-2*
C14	2.0	0.10		*5.3*	0.57		*29*	0.09		*4*
C26	1.8	0.01		*0.8*	0.02		*1*	0.13		*7*
C6	5.3	0.01		*0.1*	0.13		*3*	-0.19		*-4*
C31	0.1	-0.01		*-7.8*	-0.03		*-34*	-0.01		*-8*
C66	0.1	-0.01		*-23*	-0.02		*-32*	0.00		*-1*
C9	5.6	-0.02		*-0.3*	1.44	g	*26*	0.23	g	*4*
C2	8.4	-0.29	l	*-3.5*	-0.72	l	*-9*	-0.57	l	*-7*
	M$/day	*M$/day*		%	*M$/day*		%	*M$/day*		%

*g indicates the company that gained the most in daily profits under the scenario
*l indicates the company that lost the most in daily profits under the scenario

This analysis suggests that the distributional effects of a differentiated regulation might not differ dramatically from the application of a more stringent aggregate cap on NO_X emissions. In the Stringent Aggregate Cap scenario, the maximum amount gained by a company was about $1.5 million (79%) compared to a similar maximum gain of $1.4 million (26%) in the Differentiated Peak case and a smaller maximum gain of $0.2 million (4%) in the Differentiated Average case (indicated by "g" in Table 5-11). The maximum loss in the Stringent Aggregate Cap scenario was about $0.3 million (3%) compared to larger maximum losses of about $0.7 million (9%) and $0.6 million (6%) in the Differentiated Peak and Average cases respectively (indicated by "l" in Table 5-11). By combining the two differentiated scenarios, assuming that the NO_X reductions might be required in some peak and some average hours, the average change in profits would be about $83,000 in the Stringent Aggregate Cap scenario compared to $40,000 in the differentiated case.

The differentiated program does, however, differ in an important way from the more stringent aggregate cap. The distributional affects of the differentiated cap-and-trade program would depend on the hours in which the more stringent exchange rates applied. For example, some companies could gain a lot if the more stringent exchange rate applied in peak as opposed to average hours. Although the two companies that gained and lost the most in the two differentiated cases were the same (Table 5-11), some companies that lost in one scenario gained in the others (as also suggested by Table 5-10). For example, the Peak scenario increased the profits of C115 by about 15% but the Average scenario decreased C115's profits by about 19%. This would be an incentive for companies like C115 to try to affect the definitions of the differentiated program by offering their own models that suggested when the reductions were needed. As discussed in Section 3.3.3, if they succeeded it could impact the effectiveness of a differentiated regulation. However, in the example of the Clean Air Interstate Rule discussed in Section 3.3.3, the attempts of state governments and industry to alter the EPA's definitions did not succeed. This suggests that a differentiated regulation could also withstand this implementation challenge.

Future work will integrate this analysis with atmospheric chemistry modeling to understand which reductions would actually be necessary to reduce ozone concentrations.[155] This analysis would also help refine the analysis of distributional impacts.

5.3. Summary

The problem of the continued nonattainment of ozone air quality standards in the Eastern U.S. may lie in the mismatch between the relatively uniform incentives to reduce NO_X provided by existing regulatory systems and the highly variant temporal and locational impact of NO_X precursor emissions on ozone formation in any given area. A time- and location-differentiated cap-and-trade program implemented using ozone forecasting to alter NO_X emission permit exchange ratios in a wholesale electricity

---

[155] This work is currently being performed by the Center for Energy and Environmental Policy Research.

market that uses bid-based, security-constrained economic dispatch could help the states in the Eastern U.S. reduce the likelihood of peak ozone episodes cost effectively. One of the necessary conditions for this type of regulation is that power plants have the flexibility to reduce NO_X emissions in the short term.

Two simulation methods suggest that the potential magnitude of NO_X reductions from the redispatch of generating units in the area of Classic PJM, while taking transmission constraints into account, is between 6 tons (or 15%) on the highest demand days of 2005 in Classic PJM and 8 tons (or 30%) on average demand days. The magnitudes of potential hourly reductions depend on the time of day and the corresponding level of electricity demand. These region-wide net reductions are not accompanied by "hotspots" – large increases in NO_X in subareas of Classic PJM. In addition, redispatch is only one way that power plants can reduce emissions in the short term. Some control technologies can be used to alter emission rates on the timescale of a few weeks. In the longer term, high NO_X prices would also provide incentives for power plants to invest in NO_X control technologies.

Optimal power flow simulations of the potential NO_X reductions from redispatch were more conservative than zonal model simulations, which were based on an abstract network graph representation of the power network. The zonal model estimated the potential emissions reductions reasonably well, but allowed many more exchanges of generation between power plants than did the optimal power flow simulations. This suggests a need to account for the details of a power system's physical parameters and for network constraints other than those that are initially binding in redispatch simulations.

Although future work is needed to link the estimates of potential NO_X reductions to atmospheric chemistry modeling, the results of these simulations are encouraging. They suggest that an important pre-condition for the implementation of a time and location differentiated regulatory system is satisfied, namely, the existence of significant flexibility to reduce NO_X precursor emissions through the redispatch of power plants on hot summer days when ozone formation is most likely and the electricity system is most likely to be constrained.

Chapter 6 – Summary and Future Work

6.1. Summary of the motivation for this research

This work contributes to a larger research effort that is seeking to determine whether a NO_X cap-and-trade program that differentiated between emissions by time and location could reduce ozone concentrations more cost-effectively than reductions in the aggregate cap now in place. The difficulties that some highly populated areas of the Eastern U.S. have had attaining the federal air quality standards for ozone provides the primary motivation for this inquiry. Three factors suggest that differentiation between NO_X emissions at different times and locations might help solve this problem. First, the atmospheric chemistry literature shows that the timing and location of NO_X emission reductions determines their effectiveness in preventing peak ozone concentrations. Second, the environmental economics literature cautions that the failure to address the temporal and spatially important features of environmental problems can lead to inefficiencies because a lack of differentiation can cause costly over-compliance in some areas and times and under-compliance in others. Third, despite the atmospheric chemistry and economic justifications for doing so, the cap-and-trade programs (and other regulations) designed to address ozone pollution in the Eastern United States have not accounted for the influence of the location and timing of NO_X emissions on ozone formation.

6.1.1. Atmospheric chemistry evidence supporting differentiation

The atmospheric chemistry literature suggests that the lack of differentiation across NO_X emissions emitted at different times and locations may contribute to the inability of seasonal cap-and-trade programs in the Eastern United States to help all areas attain the ozone air quality standards. The reason for this is that the reactions governing ozone formation depend on the relative concentrations of VOCs and NO_X and on the temperature and amount of sunlight. Decreases in NO_X emissions will limit ozone formation the most during hot, sunny conditions in areas and times with high concentrations of VOCs. In areas and times of low VOC emissions and high concentrations of NO_X, a reduction in NO_X can actually *increase* ozone concentrations.

One of the clearest examples from the literature of the implications of this chemistry is a study by Mauzerall *et. al.* (2005) that examined differences in ozone formation from NO_X emissions from large point sources at different locations and times in the Eastern United States. They chose times and locations that captured relevant ranges of variation in temperature and local biogenic VOC emissions and found that the same amount of NO_X emissions at these different times and places can cause ozone formation that varies by factor of five. This study and others suggest that regulations aimed to reduce ozone concentrations should account for the impact of the timing and location of NO_X emissions on ozone formation but they do not address the details of how such a regulation could be implemented.

6.1.2. The mismatch between science and policy

Congress first recognized the environmental and health effects of ozone and its precursor emissions when it passed the Clean Air Act of 1970. Subsequent policies and regulations in the United States, and specifically in the Eastern United States, have reduced emissions of NO_X and VOCs from both mobile and stationary sources. The joint discussion of the science and policy background of the ozone problem helps explain why policies to reduce ozone concentrations have not been entirely successful even though they have caused reductions in ozone precursor emissions.

Policymakers have adjusted their approach to controlling ozone concentrations as our scientific understanding of the problem improved. For example, prior to 1990, the EPA focused on regulating VOC emissions from stationary sources in order to reduce ozone pollution. As science stressed the importance of reducing NO_X emissions, the EPA shifted their focus. However, with the exception the California Air Resources Board's recognition of the weekend effect in southern California and the rough seasonality of the ozone cap-and-trade programs in the Eastern United States, regulations to reduce ozone precursor emissions have not accounted for the impact of the timing and location of NO_X emissions on ozone formation.

6.1.3. The economic rational for differentiation

The efficiency losses from treating dissimilar emissions alike provide the basic economic rational for differentiated regulations. Regardless of the reasons that the impacts of two unit-quantities of emissions differ (e.g. location, timing, or chemical composition), treating dissimilar emissions as the same creates the need for costly over-control. For example, a un-differentiated cap-and-trade program for emission permits equalizes marginal abatement costs among sources by allowing the sources with the lowest marginal abatement cost to reduce emissions and sell excess permits to sources with higher marginal abatement costs. Under a un-differentiated program, the sources that impact air quality the most in a targeted area, like an ozone NAAQS nonattainment area, only reduce emissions if their marginal abatement costs are less than those at all other sources, including those that have little effect on the air quality in a targeted area. This means that too little is spent to control the most harmful emissions and too much is spent to control those that are least harmful.

The regulator's only control mechanism in the un-differentiated case is the overall cap on emissions. In order to meet air quality goals in stubborn areas, the regulator must lower the overall cap on emissions until the air quality is acceptable at all locations. This causes two inefficiencies: over-control at sources emitting emissions with little impact on targeted areas and the necessity of larger aggregate reductions.

6.1.4. Barriers to the implementation

There are barriers to the implementation of a differentiated cap-and-trade program for NO_X and these help explain why regulations have not addressed the temporal and spatial aspects of the ozone problem. These barriers are technical, economic, political, and legal. Technical barriers include the modeling requirements needed to support a differentiated regulation. Experience with ozone air quality forecasting and improvements in atmospheric chemistry modeling techniques may now mitigate this challenge. Ongoing research, of which this dissertation is a part, is verifying whether the modeling and forecasting requirements of a differentiated regulation are feasible.

Two other barriers to implementation are whether emission sources could respond to incentives for NO_X reductions that changed in time and by location and whether the distributional impacts of a differentiated regulation could constrain its effectiveness. This dissertation evaluated these barriers and found that although they are not trivial challenges, they could be overcome. The details of this analysis are summarized in the following sections.

6.2. Findings and contributions

6.2.1. An example benefit of a differentiated NO_X cap-and-trade program

The analysis presented in Chapter 4 suggested that seasonal, un-differentiated cap-and-trade programs in the Eastern United States did not guarantee that only the most cost-effective NO_X emission reductions occurred. This illustrated an opportunity for a differentiated cap-and-trade program to improve upon the current regulations.

Chapter 4 examined the NO_X emission characteristics of coal power plants and the hypothesis that the operators of coal power plants reduced NO_X emissions at the times when ozone formation was least likely under the summertime cap-and-trade programs in the Eastern United States in 2002 and 2005. This behavior was economically rational because it was consistent with the incentives that the operators of some coal power plants have to reduce NO_X emissions when electricity prices are low.

Three incentives were discussed that explain why it could have been less costly for the operators of some coal generating units to abate NO_X emissions when power prices were low. First, output reduction is an effective NO_X control strategy because some coal units' NO_X emission rates decrease at decreased levels of output and reduced output is less costly when power prices are low. Second, if a generating unit's marginal costs increase more steeply with a NO_X control technology than without it the operator might prefer to use the technology only during periods with low expected power prices. Third, if a NO_X control technology increases a generating unit's heat rate it can decrease its effective capacity, which is more costly when power prices are high.

The estimated NO_X abatement of specific power plants under the OTC NO_X Budget Program in 2002 and of a larger set of coal power plants in 2002 and in 2005 under the SIP NO_X Budget Program provided evidence that the abatement strategies for some coal plants were consistent with these incentives. In one example, nearly all of a generating unit's NO_X abatement occurred in the first two months of the ozone season when power prices were lower than in the latter three months. Aggregate estimates of NO_X abatement in 2002 and 2005 from coal generating units in the OTC regions of the New England and PJM power systems suggested that, although NO_X reductions occurred during the entire ozone season, slightly more abatement took place in the earlier weeks. Atmospheric chemistry modeling would be needed to determine whether the additional NO_X reductions of about 400 tons per week that occurred in the early weeks of the ozone season might have helped mitigate peak ozone concentrations if they had also occurred in the later weeks. But the analysis of the abatement decisions of the operators of these coal plants shows that a differentiated cap-and-trade program could improve the cost-effectiveness of the current cap-and-trade programs by providing stronger incentives for the emission reductions with the greatest potential to reduce ozone concentrations.

6.2.2. Flexible NO_X reductions from power plants in Classic PJM

One of the technical barriers to the implementation of a differentiated cap-and-trade program is whether or not sources could respond to incentives that change in time and by location. In the case of stationary sources in the Eastern U.S., a differentiated regulation would require the response of power plants because they contribute about 90% of NO_X emissions from stationary sources. A commonly held misperception that this work dispelled was that high levels of generation and network constraints would limit the flexibility of power plant operators to reduce emissions with redispatch in peak electricity demand hours.

Chapter 5 presented the results of the simulations that used two methods to estimate the potential short-term NO_X reductions from power plants under a differentiated cap-and-trade program. The simulations were performed for the electric power network that covers Pennsylvania, New Jersey, Maryland, Delaware, and the District of Columbia (called "Classic PJM"). Both simulation methods – a zonal model based on an abstract representation of the network and Optimal Power Flow (OPF) simulations – accounted for network constraints. Both methods suggested that power plant redispatch could reduce hourly NO_X emissions from power plants by between 6 tons (or 15%) on the highest demand days and 8 tons (or 30%) on average demand days. These potential reductions are nearly twice those being required on peak electricity demand days by a recent Memorandum of Understanding in OTC states in the Northeastern United States. Simulations of the NO_X prices required to cause redispatch also indicated that the costs of redispatch would be similar or less than those considered as options in the OTC Memorandum of Understanding.

Similar reductions to those estimated for Classic PJM would likely be available from the other Eastern U.S. power systems because they all share three key characteristics. First, the diversity of the NO_X emission characteristics of coal-fired power plants contributed to the potential NO_X reductions from redispatch in PJM (the analsyis in Chapter 4 demonstrated that the New York and New England power systems share this characteristic). Second, operating reserve margins also contributed to the flexibility to reduce NO_X emissions in Classic PJM: the other two power systems also maintain reserve margins.[156] Third, New England and New York also fill demand with a mixture of coal, oil, and gas generating units – the last key characteristics that led to the availability of flexible NO_X emission reductions from PJM.

The analysis described in Chapter 5 found that network congestion did not greatly impact the potential short-term emission reductions from redispatch in Classic PJM. The results of these simulations were a useful contribution because of the commonly held misperception that high levels of generation and network constraints would limit the flexibility to reduce emissions with redispatch in peak electricity demand hours. The finding of emissions flexibility as well as the comparison of the zonal and optimal power flow methods stressed the importance of a detailed representation of electric power systems when addressing questions about the ability of the systems to respond to new challenges, environmental or otherwise.

Section 5.2.3 offered three explanations for the finding that network constraints had little effect on emissions flexibility. The first was related to the fact that there was variation in the low and high NO_X generating units within most relatively small subsections of the interconnected network. High NO_X units are not mostly in one area of PJM and low NO_X units in another; they tend to be located together *within* the zones created by transmission constraints. This is particularly important in high demand hours

---

[156] The Federal Energy Regulatory Commission and the North American Electric Reliability Corporation require this.

because congestion is less of a problem if local demand is predominantly filled by local generation. If there is significant local NO_X-rate heterogeneity then NO_X emissions can be reduced without substantial increases in the utilization of transmission lines. The second reason for the small effect of transmission constraints was that when low-NO_X generation was located at one end of a congested line, it tended to be located on the on the high-LMP side of the constraint. This meant that increasing the output from the low-NO_X generator reduced congestion. The third and final reason for the small impact of transmission constraints was simply that the redispatch that reduced NO_X emissions involved only small amounts of generation, especially in the peak hour.

Security-constrained optimal power flow modeling was also used to examine whether a tradeoff occurred between the environment (in the form of NO_X emissions) and enforcing security constraints in Classic PJM. No clear tradeoff existed because although NO_X emissions from security-constrained dispatch simulations were about 5% higher compared to simulations without security constraints in peak demand hours, NO_X emissions were about 18% lower in security-constrained simulations in average demand hours. The addition of security constraints did not greatly change the potential NO_X reductions from redispatch in the peak demand simulations. However, they did reduce the potential reductions in average demand hours. This occurred because a larger amount of generation is available for redispatch in the average demand hours than in the peak hours and the additional constraints limited some of the potential exchanges in the average case. In the peak case, so little generation was exchanged to obtain the NO_X reductions that the security constraints did not have an effect.

The potential distributional effects of a differentiated regulation would depend greatly on the particular reductions needed to mitigate ozone episodes. Further work currently being undertaken will help determine these details by integrating atmospheric chemistry modeling with the simulations of the potential flexibility in NO_X from power plants from this work. The variation in the potential distributional effects of the redispatch simulations modeled here does suggest that industry or state governmenta would likely attempt to use political means or legal arguments to influence the details of a differentiated regulation because small changes in the determination of the timing and location of reductions needed to mitigate ozone concentrations would have large impacts on the winners and losers from the regulation.

There are two reasons that this could limit the effectiveness of a differentiated regulation more than a un-differentiated regulation. In the case of a differentiated cap-and-trade program, small changes in the definitions of trading ratios or boundaries could have large impacts on particular companies or states. This is true of any regulation. But in the differentiated case these changes could also impede the effectiveness of the differentiated regulation. As discussed in Section 3.3.3, companies typically dispute the EPA's modeling and the resulting definitions of required emission reductions by offering their own modeling results that show that they are less responsible for the air quality problems. This could more of a problem in the differentiated case if the uncertainties in the models were greater at, for example, the finer spatial resolution required to support a the differentiation. However, in the most recent example of the Clean Air Interstate Rule the disputes did not ultimately cause the EPA to adjust their initial decision. This suggests that it would be possible for a differentiated regulation to be implemented successfully. In addition, a differentiated program could reduce the total amount of NO_X abatement required by sources because it would not require emission reductions that would not help solve the ozone problem. This could reduce the overall costs to industry of the NO_X cap-and-trade programs for ozone, which could generate support for the differentiated approach.

6.3. Conclusion

The broad conclusion, or recommendation, of this dissertation is that further analysis is warranted of the topic of whether a differentiated cap-and-trade program for NO_X emissions is an implementable and cost-effective strategy to reduce ozone concentrations in the Eastern United States. The reasons for this are three-fold:

* First, ozone air quality standards have been difficult to attain in all areas and the atmospheric chemistry literature suggests that a the mismatch between policy and science is one reason for this difficulty. Existing NO_X cap-and-trade programs require stationary sources in the Eastern U.S. to reduce NO_X emissions without reference to timing or location although the timing and location of emissions influence the amount of ozone formed from a given amount of emissions. Recent advances in air quality modeling and forecasting may also mitigate implementation challenges.

* Second, differentiation could provide efficiency gains and the analysis of historical compliance data from coal power plants in this work offered an example of the possible gains relative to existing regulations.

* Third, simulations of an electric power network in the Eastern U.S. found that, given network constraints, power plants could respond to differentiated incentives with nontrivial NO_X emission reductions. The finding that network constraints did not greatly affect the flexibility for potential NO_X reductions from the redispatch of power plants allayed the common misperception that they would.

Further research is needed before the implementation of a differentiated cap-and-trade program can be fully recommended because the potential NO_X reductions from power plants must be linked to atmospheric chemistry modeling to show whether or not they could mitigate ozone formation in target areas (see Section 6.4.1). If this modeling shows that the reductions from power plants could help some areas attain the ozone air quality standards, the distributional impacts of the needed reductions could be assessed in more detail in order to understand whether a differentiated regulation could be effectively implemented. The costs of potential reductions from power plants should also be compared to those of reducing emissions further from other sources (Section 6.4.2).

This dissertation suggests that these further research challenges are worth pursuing because a differentiated regulation could help some highly populated areas of the Eastern U.S. attain the ozone air quality standards cost-effectively. If the EPA implements the more stringent ozone air quality standards it is now considering, the problem of nonattainment and the need for a cost-effective solution will become more widespread. In addition, the ozone problem provides an opportunity for regulators to experiment with a spatially and temporally differentiated cap-and-trade program. Just as experiments with cap-and-trade programs have proven fruitful, so too could experience with differentiated cap-and-trade program because ozone pollution is not the only environmental problem with spatial and temporal dimensions. For example, particulate matter pollution also has complex elements on spatial and temporal dimensions and systems to forecast particulate matter are being developed (e.g. Ojha *et al.* 2002 and Lee *et al.* 2003).

6.4. Opportunities for further work

6.4.1. Integrating potential NO_X reductions with atmospheric chemistry modeling

Work is currently being undertaken at the Center for Energy and Environmental Policy Research at MIT that will link the estimates of potential reductions from power plants described in this dissertation to weather forecasting and atmospheric chemistry models in order to determine if the simulated NO_X reductions are of the necessary magnitude to reduce the likelihood of ozone episodes. Then changes in ozone concentrations can be matched to estimates of marginal damages from the literature (e.g., Mauzerall *et. al.* 2005). This will enable evaluation of the economic opportunities to use time and locational variations in emissions prices to take advantage of the physical opportunities to reduce NO_X emissions. It will also enable a more thorough analysis of the potential distributional impacts of the specific NO_X reductions that a differentiated regulation might require.

6.4.2. Mobile sources

The focus in this work, and the primary focus of regulators, has been on reducing NO_X emissions from electric generators, but another option is to tighten controls on NO_X emissions from mobile sources. Mobile sources could also respond to incentives that changed based on forecasts. For example, the variable cost of using selective-catalytic reduction (SCR) on diesel trucks is high due to the cost of urea. The use of these controls could be mandated only in locations and at times when the NO_X reductions would reduce the formation of ozone in highly populated areas. A pricing system could also be used to deter driving during specific periods and in highly populated areas where the resulting reductions in NO_X emissions would reduce the likelihood of high ozone concentrations. Because controlling NO_X emissions from vehicles has not been thoroughly analyzed as an option to target ozone episodes, it is difficult to find cost information to compare to the above estimates of short-term reductions in NO_X from stationary sources. But because little has been done to reduce NO_X from mobile sources, especially in comparison to the number and stringency of NO_X regulations on stationary sources, it is possible that the reductions would be less expensive than further reductions from stationary sources.[157]

---

[157] In a general, non-targeted sense, the cost effectiveness of retrofitting heavy-duty on-road vehicles with SCRs is about $5,000/ton over the lifetime of the equipment. EPA, "NO_X Mobile Measures", available at

One study estimates that the margin costs of reducing NO_X emissions from mobile sources are currently about one-fifth of those from obtaining further reductions from power plants (Fowlie *et al.* 2007).

6.4.3. Comparison of theoretical network models to optimal power flow models

The abstract network model of the Classic PJM power system constructed in this dissertation provides an opportunity to examine questions posed in the literature on theoretical networks. For example, the literature has characterized power systems other than Classic PJM as having degree distributions that are power laws with an exponential cutoff.[158] This means that, up to some degree, highly connected nodes are common in the networks. These nodes are called hubs and the literature suggests they may increase the robustness of a network because a system with hubs will remain interconnected in the face of failures at the many nodes that are not hubs. The abstract network graph and optimal power flow models of Classic PJM could be used to test the predictions of network theory about whether the network could withstand the removal of nodes with actual optimal power flow simulations of the impacts of removing nodes on the network.

---

www.epa.gov/air/ozonepollution/SIPToolkit/documents/nox_mobile_measures.pdf.

[158] See, for example, Amaral, L.A.N., *et al.* (2000).

References

Adams, W. C. (2006), “Comparison of chamber 6.6-h exposures to 0.04–0.08 PPM ozone via square-wave and triangular profiles on pulmonary responses.” *Inhalation Toxicology* **18**: 127-136.

Amaral, L.A.N., *et al.* (2000), “Classes of behavior of small-world networks. *Proc. Natl. Acad. Sci. USA* **97**: 11149.

Atkinson, S.E., and T.H. Tietenberg, (1982), “The Empirical Properties of Two Classes of Designs for Transferable Discharge Permit Markets,” *Journal of Environmental Economics and Management* **9**: 101-121.

Army Corp of Engineers (Army Corp 1988), “Air Pollution Control Systems for Boilers and Incinerators,” *Departments of the Army and the Air Force*, ARMY TM 5-815-1. See Chapter 11, “Nitrogen Oxide Control and Reduction Techniques,” in particular.

Banzhaf, S., D. Burtraw, and K. Palmer, (2004), “Efficient Emission Fees in the U.S. Electric Sector,” *Resource and Energy Economics*, **26**(3), September; 317-341.

Bates, D.V. (1983), “Epidemiological Basis for Photochemical Oxidant Standard,” *Environmental Health Perspective,* **52**; 125-29.

Bell, M.L., A. McDermott, S.L. Zeger, J.M. Samet, and F. Dominici, (2004), “Ozone and Short-term Mortality in 95 US Urban Communities 1987-2000,” *J. Am. Med. Assoc.* **292**; 2372-2378.

Bell, M.L., F. Dominici, J.M. Samet, (2005), “A Meta-Analysis of Time-Series Studies of Ozone and Mortality With Comparison to the National Morbidity, Mortality, and Air Pollution Study,” *Epidemiology* **16**; 436-445.

Blanchard, C.L. and S. Tanenbaum, (2003a), “Trends in Ambient NO_X and Particulate Nitrate Concentrations in California, 1980-2000, ” *Final Report, Coordinating Research Council,* CRC Project No. A-43c.

Blanchard, C.L. and S. Tanenbaum, (2003b), “Differences between weekday and weekend air pollutant levels in southern California, ” *J. Air Waste Manage Assoc.* **53**(7); 816-28.

Burtraw, D., K. Palmer, R. Bharvirkar, and A. Paul, (2001), “Cost-Effective Reduction of NO_X Emissions from Electricity Generation,” *J. Air Waste Manage Assoc* **51**;1476–89.

Burtraw D., R. Bharvirkar, M. McGuinness, (2003), “Uncertainty and the cost-effectiveness of NO_X emission reductions from electricity generation,” *Land Economics* **79**(3): 382-401.

Burtraw D., and D.A. Evans, (2004), "NO_X emissions in the United States: a potpourri of policies," In *Choosing Environmental Policy: Comparing Instruments and Outcomes in the United States and Europe.* Winston Harrington, Richard D. Morgenstern and Thomas Sterner, Eds. Resources for the Future: Washington D.C.

Burtraw, D., D.A. Evans, A. Krupnick, K. Palmer, and R. Toth, (2005), "Economics of Pollution Trading for SO_2 and NO_X," *Resources for the Future Discussion Paper,* March.

Bond, D.W., R. Zhang, X. Tie, G. Brasseur, G. Huffines, R.E. Orville, and D.J.Boccippio, (2001), "NO_X production by lightning over the continental United States," *Journal of Geophysical Research*, **106**(D21); 27,701-710.

Borrell, P., (2003), "The Weekend Effect," *The Royal Society of Chemistry*: *chembytes e-zine,* May (http://www.chemsoc.org/chembytes/ezine/2003/borrell_may03.htm).

CARB: California Air Resources Board, (2004), "The Ozone Weekend Effect in California" at http://www.arb.ca.gov/aqd/weekendeffect/weekendeffect.htm retrieved November 15, 2006.

Chameides, W.L., R.W. Lindsay, J. Richardson, and C.S. Kiang, (1988), "The Role of Biogenic Hydrocarbons in Urban Photochemical Smog: Atlanta as a Case Study," *Science* **241**(4872); 1473-1475.

Chen, W.Y. (1989), "Estimate of dynamical predictability from NMC DERF Experiments," *Monthly Weather Review,* **117**; 1227-1236.

Chevron, (2007), "Motor Gasolines Technical Review", San Ramon, CA, at http://www.chevron.com/products/prodserv/fuels/bulletin/motorgas/downloads/.

Crocker, T.D. (1966), "The Structuring of Atmospheric Pollution Control Systems," in *The Economics of Air Pollution*, (Harold Wolozin, Ed.), New York: W.W. Norton & Company, Inc.

Dales, John, (1968), *Pollution, Property and Prices*, Toronto: University Press.

Davidson, P.M., N. Seaman, K. Schere, R.A. Wayland, J.L. Hayes, and K.F. Carey, (2004), "National Air Quality Forecasting Capability: First Steps toward Implementation," presented at the Sixth Conference on Atmospheric Chemistry: Air Quality in Megacities Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone, January.

Decker, C.E., L.A. Ripperton, J.J.B. Worth, F.M. Vukovich, W.D. Bach, J.B.-Tommerdahl, F. Smith, and D.E. Wagoner (1976). "Formation and Transport of Oxidants Along Gulf Coast and in Northern U.S." Research Triangle Park, NC: U.S. Environmental Protection Agency, *EPA-450/3-76-003.*

Eder, B. and S. Yu, (2006), "A performance evaluation of the 2004 release of Models-3 CMAQ," *Atmospheric Environment* **40**(26); 4811-4824.

Ellerman, A.D., R. Schmalensee, P.L. Joskow, J.P. Montero, and E.M. Bailey, (1997), *Emissions Trading under the U.S. Acid Rain Program*, Cambridge, MA: Center for Energy and Environmental Policy Research, Massachusetts Institute of Technology.

Ellerman, A.D. (2000), "From Autarkic to Market-Based Compliance: Learning from Our Mistakes," in *Emissions Trading,* ed. R.F. Kosobud, New York: John Wiley & Sons, Inc.

Ellerman, A.D., P.L. Joskow, R. Schmalensee, J.P. Montero, and E.M. Bailey, (2000), *Markets for Clean Air: The U.S. Acid Rain Program,* New York, NY: Cambridge University Press.

Ellerman, A.D., (2004), "The U.S. SO2 Cap-and-Trade Programme," in *Tradeable Permits: Policy Evaluation, Design and Reform,* Organization for Economic Co-Operation and Development.

Farrell, A.E., R. Carter, and R. Raufer, (1999), "The NO_X budget: market-based control of tropospheric ozone in the northeastern United States," *Resource and Energy Economics* **21**(1): 103-24.

Farrell, A.E. (2003) "Temporal Hotspots in Emissions Trading Programs: Evidence from the Ozone Transport Commission NO_X Budget," Presented at an EPA conference, *Market Mechanisms and Incentives: Applications to Environmental Policy* (Washington, D.C., May 2003).

Fiore, A.M., D.J. Jacob, I. Bey, R.M. Yantosca, B.D. Field, and A.C. Fusco, (2002), "Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes," *Journal of Geophysical Research* **107**(D15); 10.1029.

Fiore, A.M., L.W. Horowitz, D.W. Purves, H. Levy II, M.J. Evans, Y. Wang, Q. Li, and R.M. Yantosca, (2005), "Evaluating the contribution of changes in isoprene emissions to surface ozone trends over the eastern United States," *Journal of Geophysical Research*, **110**(D12303), doi:10.1029/2004JD005485.

Fowlie, M., (2005), "Emissions Trading, Electricity Industry Restructuring, and Investment in Pollution Abatement," Center for the Study of Energy Markets Working Paper, CSEM WP 149, University of California at Berkeley.

Fowlie, M., C.R. Knittel, and C. Wolfram, (2007), "Sacred Cars: Optimal Regulation of Stationary and Non-stationary Pollution Sources," University of California Berkeley, Haas School of Business Working Paper, available at http://www.econ.ucdavis.edu/faculty/knittel/papers/NOx.pdf

Harrison, D. Jr., (2004), "Ex-post evaluation of the RECLAIM Emissions Trading Programmes for the Los Angeles Air Basin," in *Tradeable Permits: Policy Evaluation, Design and Reform,* Organization for Economic Co-Operation and Development.

Hogan, W., (1998), "Competitive Electricity Market Design: A Wholesale Primer," Harvard University, John F. Kennedy School of Government working paper.

Hogan, W., (1999), "Transmission Congestion: The Nodal-Zonal Debate Revisited," Harvard University, John F. Kennedy School of Government, available at http://ksghome.harvard.edu/~whogan/nezn0227.pdf, retrieved November 21, 2006.

ICF Resources, Inc., (1989), "Economic, Environmental, and Coal Market Impacts of SO_2 Emissions Trading Under Alternative Acid Rain Control Proposals," Regulatory Innovations Staff, USEPA.

Ito, K.; S.F. De Leon, M. Lippman, (2005), "Associations Between Ozone and Daily Mortality: Analysis and Meta-Analysis," *Epidemiology* **16**; 446-457.

Lamont, J.W. and E.V. Obessis (1995), "Emission dispatch models and algorithms for the 1990s," *IEEE Trans. Power Syst.* **10**(2); 941–947.

Lee, Y.D., S. Yun, and Y. Lee (2003), "Analyzing weather effects on airborne particulate matter with HGLM," *Envionmetrics* **14**(7); 687-97.

Levy, J. I., S.M. Chemerynski, and J.A. Sarnat, (2005), "Ozone Exposure and Mortality: An Empiric Bayes Metaregression Analysis," *Epidemiology* **16**: 458-468.

Joskow, P.L., and R. Schmalensee, (1998), "The Political Economy of Market-Based Environmental Policy: The U.S. Acid Rain Program," *The Journal of Law & Economics*, **41**(1); 37-83.

Joskow, P.L. (2006), "Markets for Power in the United States: An Interim Assessment," *Energy Journal*, **27**(1); 1-36.

Kang, D., V. P. Aneja, R. Mathur, and J. D. Ray, (2003), "Nonmethane hydrocarbons and ozone in three rural southeast United States national parks: A model sensitivity analysis and comparison to measurements," *Journal of Geophysical Research*, **108**(D19): 4604, doi:10.1029/2002JD003054.

Kang, D., B.K. Eder, A.F.Stein, G.A. Grell, S.E. Peckham, and J. McHenry, (2005), "The New England Air Quality Forecasting Pilot Program: Development of an Evaluation Protocol and Performance Benchmark," *Air and Waste Management,* **55**; 1782-96.

Klier, T.H., R.H. Mattoon, and M.A. Prager, (1997), "A mixed bag: assessment of market performance and firm trading behavior in the NO_X RECLAIM Programme," *Journal of Environmental Planning and Management* **40**(6): 751-74.

Krupnick, A.J., W.E. Oates, and E. Van De Verg, (1983), "On Marketable Air-Pollution Permits: The Case for a System of Pollution Offsets," *Journal of Environmental Economics and Management* **10**: 233-247.

Krupnick, A.J., V. McConnell, M. Cannon, T. Stoessell, and M. Batz, (2000), "Cost-Effective NO_X Control in the Eastern United States," *Discussion Paper 00-18*, Washington, DC: Resources for the Future.

Lamb, B., A. Guenther, D. Gay, and H. Westberg, (1987), "A national inventory of biogenic hydrocarbon emissions" *Atmospheric Environment* **21**: 1695-1705.

Lamont, J.W. and E.V. Obessis (1995), "Emission dispatch models and algorithms for the 1990s," *IEEE Trans. Power Syst.* **10**(2); 941–947.

Lehman, J., K. Swinton, S. Bortnick, C. Hamilton, E. Baldridge, B. Ender, and B. Cox, (2004), "Spatio-temporal characterization of tropospheric ozone across the eastern United States," *Atmospheric Environment,* **38**; 4357-4369.

Lutter, R., (1999), "Is EPA's Ozone Standard Feasible?" *AEI-Brookings Joint Center for Regulatory Studies*, Regulatory Analysis 99-6, December.

Mansur, E. (2006a). "Prices vs. Quantities: Environmental Regulation and Imperfect Competition", mimeo, July, available at http://www.som.yale.edu/faculty/etm7/ retrieved November 27, 2006.

Mansur, E. (2006b). "Do Oligopolists Pollute Less? Evidence from a Restructured Electricity Market," June 8, *Journal of Industrial Economics*, forthcoming, available at http://www.som.yale.edu/faculty/etm7/ retrieved November 27, 2006.

Mauzerall, D.L., B. Sultan, N. Kim and D.F. Bradford, (2005), "NO_X emissions from large point sources: variability in ozone production, resulting health damages and economic costs" *Atmospheric Environment,* **39**(2005); 2851-2866.

McGartland, A.M. and W.E. Oates, (1985), "Marketable Permits for the Prevention of Environmental Deterioration," *Journal of Environmental Economics and Management* **12**: 207-228.

McGartland, A. M., (1984), "Marketable Permit Systems for Air Pollution Control: An Empirical Study", Unpublished Dissertation, University of Maryland, cited in Tietenburg (1995).

McKeen, S. A., E.Y. Hsie, and S. C. Liu, (1991), "A study of the dependence of rural ozone on ozone precursors in the eastern United States," *Journal of Geophysical Research,* **96**: 15,377-15,394.

Mendelsohn, R. (1986). "Regulating Heterogeneous Emissions," *Journal of Environmental Economics and Management,* **13**(4); 301-312.

Montgomery, D.W. (1972). "Markets in Licenses and Efficient Pollution Control Programs," *Journal of Economic Theory,* **5**; 395-418.

National Research Council, (NRC 1991), "Rethinking the Ozone Problem in Urban and Regional Air Pollution," Washington, DC: The National Academies Press.

Ojha, S., J. Coutinho, and A. Kumar (2002), "Developing Systems to Forecast Ozone and Particulate Matter Levels," *Environmental Progress* **21**(2); J7-12.

O'Neill, S.M., B.K. Lamb, J. Chen, C. Claiborn, D. Finn, S. Otterson, C. Figueroa, C. Bowman, M. Boyer, R. Wilson, J. Arnold, S. Aalbers, J. Stocum, C. Swab, M. Stoll, M. DuBois, and M. Anderson, (2006), "Modeling Ozone and Aerosol Formation and Transport in the Pacific Northwest with the Community Multi-Scale Air Quality (CMAQ) Modeling System", *Environ. Sci. Technol.* **40**(4); 1286 – 1299.

——— (NRC 1999), "Ozone-forming Potential of Reformulated Gasoline," Washington, DC: The National Academies Press.

——— (NRC 2001), "Evaluating Vehicle Emissions Inspection and Maintenance Programs," Washington, DC: National Academy Press.

——— (NRC 2004), "Air Quality Management in the United States," Washington, DC: The National Academies Press.

——— (NRC 2006), "State and Federal Standards for Mobile Source Emissions," Washington, DC: The National Academies Press.

Northeastern States for Coordinated Air Use Management, (NESCAUM 2006), "High Electric Demand Day and Air Quality in the Northeast," *White Paper,* June 5, available at http://www.nescaum.org/documents/high-electric-demand-day-and-air-quality-in-the-northeast/.

Organization for Economic Co-Operation and Development, (OECD 2004), *Tradeable Permits: Policy Evaluation, Design, and Reform,* OECD.

Ozone Transport Commission and the U.S. Environmental Protection Agency (OTC 2003) "NOx Budget Program: 1999-2002 Progress Report," available at http://www.epa.gov/airmarkets/otc/otcreport.pdf.

Pierce, T., C. Geron, and L. Bender, (1998), "Influence of increased isoprene emissions on regional ozone modeling," *Journal of Geophysical Research* **103**(D19): 25,611-25,629.

PJM Interconnect (PJM 2006), "2005 State of the Market Report," March 8 (Available at http://www.pjm.com/markets/market-monitor/som.html)

PJM Interconnect, (PJM 2007), "2006 State of the Market Report," (Available at http://www.pjm.com/markets/market-monitor/som.html)

Pun B.K. and C. Seigneur, (2001), "Sensitivity of particulate matter nitrate formation to precursor emissions in the California San Joaquin Valley," *Environ. Sci. Tech.* **35**: 2979-2987.

Purves, D.W., J.P. Caspersen, P.R. Moorcroft, G.C. Hurtt, and S.W. Pacala, (2004), "Human-induced changes in US biogenic volatile organic compound emissions: evidence from long-term forest inventory data," *Global Change Biology*, **10**: 1737-1755, doi: 10.1111/j.1365-2486.2004.00844.x.

Research Triangle Institute, (RTI 1975), "Investigation of Rural Oxidant Levels as Related to Urban Hydrocarbon Control Strategies." *EPA-450/3-75-036.* U.S. Environmental Protection Agency, Research Triangle Park, N.C.

Roach, F., C. Kolstad, A.V. Tobin, and M. William, (1981), "Alternative Air Quality Policy Options in the Four Corners Region," *The Southwestern Review of Management and Economics,* **1**(2): 29-58.

Ryerson, T.B., M. Trainer, J.S. Holloway, D.D. Parrish, L.G. Huey, D.T. Sueper, G.J. Frost, S.G. Donnelly, S. Schauffler, E.L. Atlas, W.C. Kuster, P.D. Goldan, G. Hu, J.F. Meagher, and F.C. Fehsenfeld, (2001), "Observations of Ozone Formation in Power Plant Plumes and Implications for Ozone Control Strategies," *Science* **292**(5517): 719-723.

Ryther, J.H. and W.M. Dunstan, (1971), "Nitrogen, Phosphorus, and Eutrophication in the Coastal Marine Environment," *Science,* **171**(3975): 1008-1013.

Sillman, S., J.A. Logan, and S.C. Wofsy, (1990), "The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes," *Journal of Geophysical Research* **95**: 1837–1851.

Smith, A.E., D.W. North, J.P. Bekemeier, N.Y. Chan, R. Glasgal, J.L. Welsh, and K.P. Green, (1997), "Costs, Economic Impacts, and Benefits of EPA's Ozone and Particulate Standards" *Reason Public Policy Institute,* Policy Study No. 226, available at http://www.rppi.org/es226.html.

Stavins, R., (1995), "Transactions Costs and Tradeable Permits" *Journal of Environmental Economics and Management* **29**; 133-148.

Stavins, R.N., (1998), "What Can We Learn from the Grand Policy Experiment? Lessons from SO_2 Allowances Trading." *Journal of Economic Perspectives* **12**(3): 69-88.

Stoft, S., (1997), "Transmission Pricing in Zones: Simple or Complex?" *The Electricity Journal* **10**(1), January/February; 24-31.

Stoft, S., (2002), *Power System Economics: Designing Markets for Electricity,* IEEE Press.

Sun, D.I., B. Ashley, B. Brewer, A. Hughes, and W. F. Tinney, (1984). "Optimal Power Flow by Newton Approach," *IEEE Transactions on Power Apparatus and Systems*, **PAS-103**, October; 2864-80.

Swift, B., (2004), "Emissions trading and hot spots: a review of the major programs," *Environmental Reporter* **35**(19): 1-18.

Tingey, D.T., M. Manning, L.C. Grothaus, and W.F. Burns, (1979), "The influence of light and temperature on isoprene emission rates from live oak," *Physiologia Plantarum* **47**:112-118.

Tingey, D.T., M. Manning, L.C. Grothaus, and W.F. Burns, (1980), "Influence of light and temperature on monoterpene emission rates from slash pine," *Physiologia Plantarum* **65**:797-801.

Tietenberg, T., (1990), "Economic Instruments for Environmental Regulation," *Oxford Review of Economic Policy,* **6**(1): 17-33.

Tietenberg, T., (1995), "Tradeable Permits for Pollution Control when Emission Location Matters: What have We Learned?" *Environmental and Resource Economics* **5**: 95-113.

Tietenberg, T. and N. Johnstone (2004), "Ex Post Evaluation of Tradeable Permits: Methodological Issues and Literature Review," in *Tradeable Permits: Policy Evaluation, Design and Reform,* Organization for Economic Co-Operation and Development.

Tietenberg, T. (2006), *Emissions Trading: Principles and Practice*, Second Edition, Washington D.C.: Resources for the Future Press.

Tong, D.Q, N.Z. Muller, D.L. Mauzerall, and R.O. Mendelsohn (2006), "Integrated Assessment of the Spatial Variability of Ozone Impacts from Emissions of Nitrogen Oxides" *Environmental Science & Technology* **40**(5);1395-1400.

Trainer, M., E. J. Williams, D. D. Parrish, M. P. Buhr, E. J. Allwine, H. H. Westberg, F. C. Fehsenfeld and S. C. Liu, (1987), "Models and observations of the impact of natural hydrocarbons on rural ozone," *Nature* **329**: 705 – 707, doi:10.1038/329705a0.

U.S. Army Corps of Engineers, (U.S. Army Corp 1988), "Air Pollution Control Systems for Boilers and Incinerators", *U.S. Army Corps of Engineers Technical Manual*, TM 5-815-1/AFR 19-6, May 9. See, Chapter 11, "Nitrogen Oxides (NO_X) Control and Reduction Techniques," at www.wbdg.org/ccb/ARMYCOE/COETM/tm_5_815_1.pdf.

U.S. Energy Information Association (U.S. EIA 2006a), "Emissions of Greenhouse Gases in the United States 2005," Department of Energy/EIA report no. DOE/EIA-0573(2005), November, available at http://www.eia.doe.gov/oiaf/1605/ggrpt/index.html.

——— (U.S. EIA 2006b), "Assumptions to the Annual Energy Outlook," DOE/EIA-0554, at http://www.eia.doe.gov/cneaf/electricity/epm/epm_ex_bkis.html

U.S. Environmental Protection Agency (U.S. EPA 1991), "Sourcebook: NO_X Control Technology Data," Offices of Research and Development, Washington DC, EPA/600/2-91/029, July.

——— (U.S. EPA 1994), "Alternative Control Technologies Document: NO_X Emissions from Utility Boilers," Emissions Standards Division, Office of Air and Radiation: Research Triangle Park, North Carolina, EPA-453/R-94-023.

——— (U.S. EPA 1998a), "Finding of Significant Contribution and Rulemaking for Certain States in the Ozone Transport Assessment Group Region for Purposes of Reducing Regional Transport of Ozone," *Federal Register*, **63**(207): 57356, October 27.

——— (U.S. EPA 1998b), "Supplemental Ozone Transport Rulemaking Regulatory Analysis," Washington, DC: Office of Air and Radiation, April.

——— (U.S. EPA. 1998c), "Regulatory Impact Analysis for the NO_X SIP Call, FIP, and Section 126 Petition, Volume 2: Health and Welfare Benefits," Washington, DC: EPA Office of Air and Radiation, December.

——— (U.S. EPA 1999), "Achieving Clean Air and Clean Water: The Report of the Blue Ribbon Panel on Oxygenates in Gasoline," *EPA420-R-99-021*, September, available at http://www.epa.gov/otaq/consumer/fuels/oxypanel/r99021.pdf.

——— (U.S. EPA 2000a), "Control of Air Pollution From New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements; Final Rule," *Federal Register*, **65**(28): 6698, February 10.

——— (U.S. EPA 2000b), "Final Emission Standards for 2004 and Later Model Year Highway Heavy-Duty Vehicles and Engines," *EPA420-F-00-026,* July, available at http://www.epa.gov/otaq/regs/hd-hwy/2000frm/f00026.pdf.

——— (U.S. EPA 2003a), "Latest Findings on National Air Quality: 2002 Status and Trends," Research Triangle Park, NC: Office of Air Quality Planning and Standards, Emissions Monitoring and Analysis Division; *EPA 454/ K-03-001*, August.

——— (U.S. EPA 2003b), "Reducing Air Pollution from Nonroad Engines," Office of Transportation and Air Quality; *EPA420-F-03-011*, April, available at http://www.epa.gov/OMS/cleaner-nonroad/f03011.pdf.

——— (U.S. EPA 2004), "Air Quality Criteria for Particulate Matter," Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office; *EPA/600/P-99/002aF*, October.

——— (U.S. EPA 2005), "Evaluating Ozone Control Programs in the Eastern United States: Focus on the NO_X Budget Trading Program, 2004," *EPA454-K-05-001*, August.

——— (U.S. EPA 2006a), "Air Quality Criteria for Ozone and Related Photochemical Oxidants," Research Triangle Park, NC: Office of Health and Environmental

Assessment, Environmental Criteria and Assessment Office; EPA report no. 600/R-05/004aF, February.

——— (U.S. EPA 2006b), "NO_X Budget Trading Program: 2005 Program Compliance and Environmental Results," EPA report no. EPA430-R-06-013, September.

——— (U.S. EPA 2007a), "Achievements in Stratospheric Ozone Protection," Office of Air and Radiation: EPA report no. EPA-430-R-07-001, April.

——— (U.S. EPA 2007b), "Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information, OAQPS Staff Paper," Research Triangle Park, NC: Office of Air Quality Planning and Standards; EPA report no. EPA-452/R-07-003, January.

van den Dool, H.M., and S. Saha (1990). "Frequency dependence in forecast skill," *Monthly Weather Review,* **118**; 128-137.

Zhang, R., X. Tie, and D.W. Bond, "Impacts of anthropogenic and natural NO_X sources over the U.S. on tropospheric chemistry," *Proceedings of the National Academy of Sciences,* **100**(4); 1505-1509.

Appendix A: Non-NO_X Price Incentives for more generation in the late ozone season

The power plants in the OTC and SIP NO_X Budget Programs participate in one of three power systems: the PJM Interconnect, the New York Power Pool, or the New England Power Pool. Although these power systems have unique characteristics, they are all based on the concept of economic dispatch. Generating unit owners bid cost curves to the system operator, who then uses these curves to dispatch the units to fill demand while minimizing total system operating costs given network constraints. Electricity prices reflect the cost to supply the system with one more megawatt of power; in a competitive market they reflect the marginal cost of the marginal generator. The system operator calls on generating units with bids lower than the price of power to generate; again, in a competitive market bids reflect units' marginal costs of providing power.

Average and maximum locational marginal prices (LMP) in 2002 and 2005 for the New England Power Pool and PJM for the "early" and "late" ozone season and the "non-ozone" season months are shown in

Table A-6-1.[159] Prices are notably higher in 2005 compared to 2002 because natural gas prices were higher (e.g. PJM 2007). The average price in the late ozone season is higher than that in the early ozone season for both regions and in both years. The maximum prices were observed in the late summer except in 2005 for New England when the peak price occurred in the winter.

---

[159] Data from the New York Power Pool are not available.

Table A-6-1 Average and maximum locational marginal prices in New England and PJM in 2002 and 2005.

LMP ($/MWh)		2002		2005	
		Average	**Maximum**	**Average**	**Maximum**
	Non-ozone	35	198	76	856
New England	**Early**	31	147	60	197
	Late	40	1000	89	341
	Non-ozone	26	137	55	262
PJM	**Early**	25	147	46	192
	Late	36	792	72	287

The duration of high prices was also longer in the late ozone season for both power systems. For example, in 2005 for PJM, prices were over $100/MWh in 30% of the late summer hours compared to 7% of the early summer hours and 10% of the non-ozone season hours (Figure A-1).

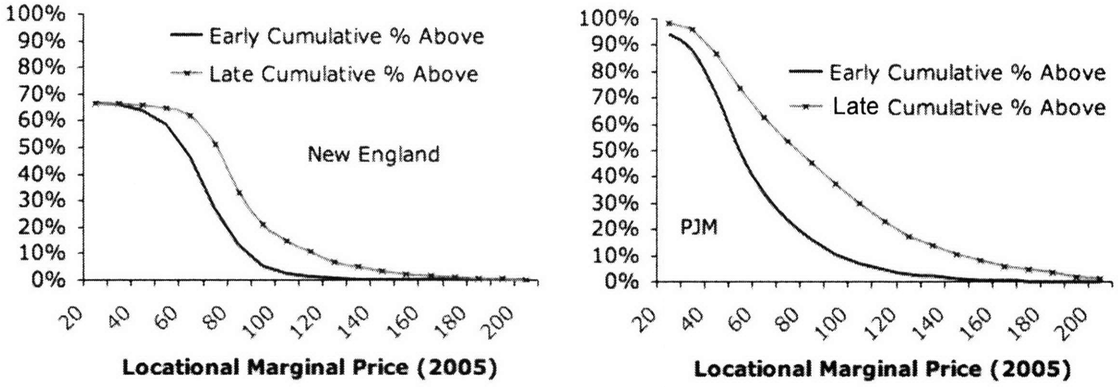

Figure A-6-1 Price duration curves for the New England Power Pool and PJM Interconnect in 2005.

Given experience with these power systems and the relationships between weather, electricity demand, and price, plant operators likely expect power prices to be higher in the late summer.

Some generators also shut down to tune NO_X control technologies, or to make combustion modifications that enable them to operate at low-NO_X firing conditions. The performance of the NO_X control strategies can then degrade over the course of the summer and the high late summer prices mean that the operators do not want to shut down the unit in order to re-tune the NO_X controls.[160]

The forced outage rate of generators can also increase when they run at full capacity for extended periods. This may be another incentive for them to operate at a lower level when prices are lower in the early summer – to ensure that the generator will be available when power prices are high. Figure A-6-2 shows the percent of unit-days each week that coal-fired generators spent entirely off during 1998 and 2002 in the OTC. This figure suggests that high electricity prices in the late summer created incentives for generators to operate more in the late summer both prior to (1998) and while (2002) the OTC NO_X Budget Program was in effect.

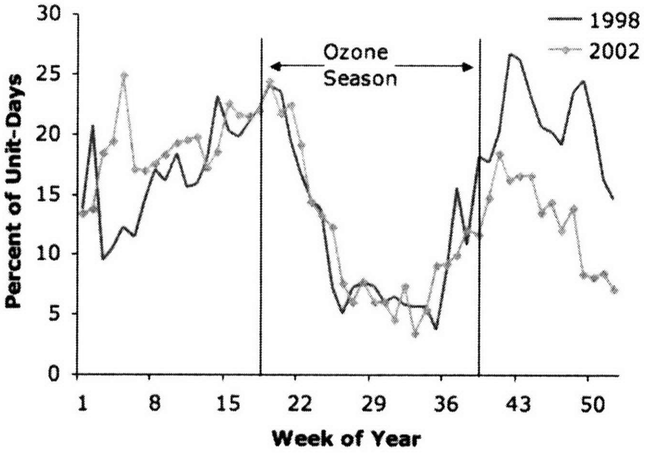

Figure A-6-2 Percentage of unit-days that coal-fired generators did not operate during 1998 and 2002.

The generators spent a higher percentage of days operating during the late ozone season in both years. The percentage of days that units did not operate started to fall from about 20% to about 5% between the beginning of June and the beginning of July and then increased again starting in September. The units spent slightly more time off in April and May than in March and previous months (with the exception of the first week of February

---

[160] Personal conversation with a former power plant operator in New England.

in 2002). These data do not suggest that generators were turned off in the early summer in 2002 as a response to the cap-and-trade program, but give an example of how high electricity demand and prices in the late summer are incentives for generators to operate as much as possible. This leads to more emissions during the months of July and August when ozone formation is the biggest problem.

Appendix B: Curve Fitting to Estimate Coal Plant Marginal Cost Curves

This appendix discusses the curve fitting results for the input-output and NO_X-output curves for the three example units. These curves were used to obtain the estimates of marginal cost curves shown in Figure 4-9. They were also used to calculate the NO_X prices required to cause the units to begin to decrease output given the electricity prices observed in 2002 (Section 4.5.3).

The input-output curves were constructed by fitting linear curves to the hourly data using ordinary least squares. A dummy variable for the ozone season months was also included. The hourly data are from the EPA's Continuous Emissions Monitoring System (CEMS). For each unit, the hourly data were used to estimate the coefficients of the relationship:

$$HI = \beta_o + \beta_1 q + \beta_2 OZONE + \beta_3 (q^* OZONE) + \varepsilon.$$

Where HI is the hourly heat input (mmBTU), *q* is the unit's hourly gross output (MW), OZONE is a dummy variable that equals one during the ozone season, and ε is the error term.

The NO_X-output curves were constructed similarly, but the data were transformed to obtain an exponential fit. For each unit, the hourly data were used to estimate the following relationship, where $\ln(NO_X)$ is the natural logarithm of the unit's hourly NO_X emissions:

$$\ln(NO_X) = \alpha_o + \alpha_1 q + \alpha_2 OZONE + \alpha_3 (q^* OZONE) + \varepsilon.$$

Table B-6-2 shows the coefficients from these regressions.

Table B-6-2 Regression results for Input-output curves (Heat Input "HI" on generation "q") and NO_X-output curves (the natural logarithm of NO_X emissions "$\ln(NO_X)$" on generation).

		OLS Coefficient (standard error in parenthesis)				**Number of Observations**	**Adjusted R^2**
		Intercept	**q**	**OZONE**	**q*OZONE**		
6094U3 2002	**HI**	110.19** (34.56)	8.73** (0.0426)	-267.0** (49.36)	0.10* (0.0609)	7260	0.921
	ln(NOx)	5.61** (0.0214)	0.003** (0.0000264)	-0.097** (0.0306)	1.77E-05 (0.0000377)	7260	0.784
		Intercept	**q**	**OZONE**	**q*OZONE**		
1571U1 2002	**HI**	-422.3** (15.19)	10.65** (0.049)	-83.64** (24.7)	0.83** (0.082)	8431	0.903
	ln(NOx)	5.40** (0.0155)	0.0065** (0.0000509)	0.20** (0.0252)	-0.00078** (0.0000833)	8431	0.740
		Intercept	**q**	**OZONE**	**q*OZONE**		
1573U2 2002	**HI**	261.0** (6.12)	8.96** (0.0011)	76.23** (7.89)	0.094** (0.0146)	6727	0.996
	ln(NOx)	6.45** (0.0115)	0.0026** (2.071E-05)	-0.17** (0.0148)	0.00024** (2.739E-05)	6727	0.864

** indicates significance at the 5% level
* indicates significance at the 10% level

Most of the coefficients are significant at the 5% level. The positive sign on the coefficients of the interaction terms for the input-output regressions (q*OZONE) indicates that the unit's heat rates were slightly higher during the ozone season as expected. The magnitudes of the estimated heat rates (coefficients on *q* in the input-output regressions) are of the right magnitude.

The results of the regressions of $\ln(NO_X)$ on output are difficult to interpret without graphs. Figure B-6-3 shows the predicted ozone season and non-ozone season NO_X-output relationship for the three plants in 2002 (calculated with the coefficients in Table B-6-2).

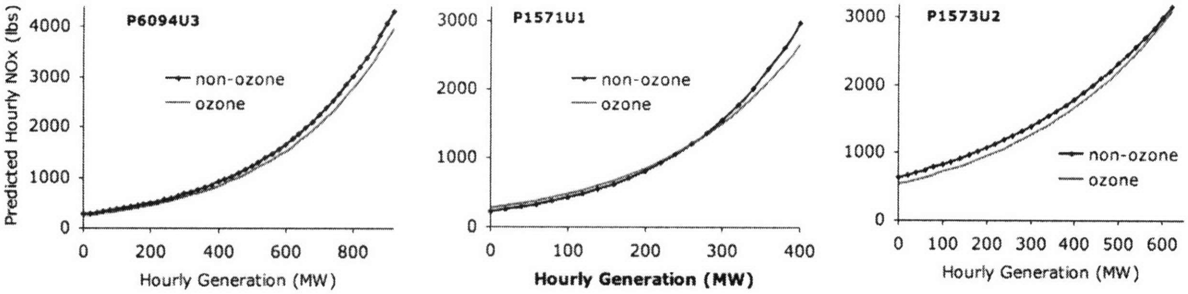

Figure B-6-3 Predicted NO_X emission versus output curves for the three example units in 2002.

The exponential fit was chosen for the NO_X-output relationship because residual plots indicated that it was superior to either a linear fit or a quadratic fit. Example plots for P1573U2 are shown in Figure B-6-4. The residuals of the exponential fit were much closer to zero for all values of output.

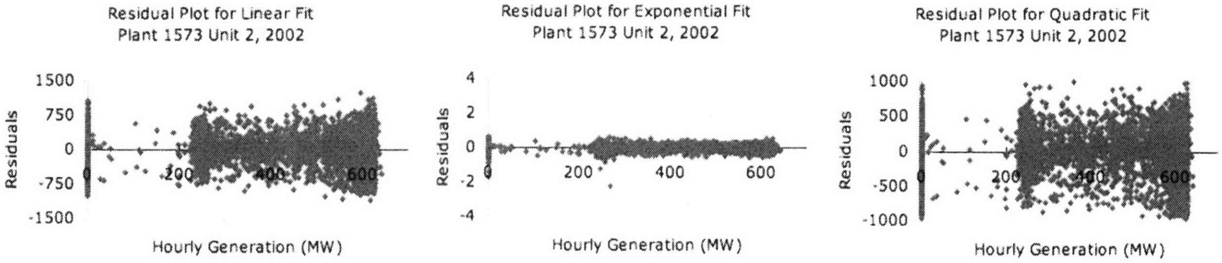

Figure B-6-4 Residual plots for regressions of linear (left), exponential (middle), and quadratic (right) fits for hourly NO_X emissions on output for Plant 1573 Unit 2 in 2002. All regressions include dummy variables for the ozone season and interaction terms for the dummy variable with output.

Appendix C: Aggregate Estimates of Generation and Emissions

This appendix presents the results of the aggregate estimates of counterfactual emissions that used electricity prices to predict generation from the same sample of coal power plants discussed in Section 4.5.4. These estimates used the same procedure as that discussed in Section 4.5.4 and the only difference was the use of electricity prices to predict generation in the first stage. To predict generation, the coefficients in the following equation were estimated:

$$q_{it} = \beta_o + \beta_{1i}\mathrm{PRICE}_t + \beta_{3i}\mathrm{EARLY} + \beta_{4i}\mathrm{LATE} + \varepsilon \qquad \text{(C-1)}$$

where PRICE was the weekly average electricity price in the power system corresponding to each generating unit (i.e. PJM or New England), EARLY was a dummy variables for May and June and LATE for July through August. The omitted variables version of the Hausman test indicated that PRICE was contemporaneously correlated with the error term in this equation. This was an expected possibility because the amounts particular units generate can affect electricity prices. Even in the case when none of the units in question set the electricity price, if coal power plants reduced their output, more generation would be required from other units with higher costs and this would increase electricity prices. To obtain consistent estimates of β_1 in Equation C-1, total weekly electricity demand was used as an instrumental variable for the average weekly electricity price. Electricity demand is a suitable instrumental variable for prices because it is not correlated with the error term (since the generation of individual units does not affect total electricity demand) and it is correlated with electricity prices.

Weekly heat input was then regressed on the predicted generation ($\tilde{q}_{it}$) from Equation C-1 and on the dummy variables:

$$HI_{it} = \alpha_o + \alpha_{1i}\tilde{q}_{it} + \alpha_{2i}\mathrm{EARLY} + \alpha_{3i}\mathrm{LATE} + \varepsilon \qquad \text{(C-2)}$$

In the last stage, weekly NO_X emissions were regressed on the predicted heat input ($\widetilde{HI}_{it}$) from Equation C-2, the square of the predicted heat input and the dummy variables:

$$NO_{Xit} = \mu_o + \mu_{1i}\widetilde{HI}_{it} + \mu_{2i}(\widetilde{HI}_{it})^2 + \mu_{3i}\mathrm{EARLY} + \alpha_{4i}\mathrm{LATE} + \varepsilon \qquad \text{(C-3).}$$

The coefficients estimated in these regressions are shown in Table C-6-3. The signs of the coefficients are as expected. For example, the sign of the coefficient on electricity prices is positive indicating that the units generated more when prices were higher. The signs on the coefficients of EARLY and LATE were negative and significant in the regressions of NO_X on predicted heat input in all years and these coefficients were larger in 2002 and 2005.

The counterfactual estimates generated from the non-ozone coefficients of these regressions suggested that in 2002 even more reductions occurred in May and June compared to July through September than the estimates in Section 4.5.4. The estimated average weekly NO_X reductions in the May and June in this case were about 1900 tons compared to about 1200 tons between July and September. In 2005 these estimates also suggested that more NO_X reductions occurred during May and June than in the later periods. However, this procedure did not predict generation and NO_X emissions as well as the procedure discussed in Section 4.5.4 (see Figure C-6-5). In particular, the predictions of NO_X emissions in 2005 were not even close. A possible reason for this is that individual generating units respond to different locational marginal prices. The average electricity price, which is the load weighted average of all the locational marginal prices, is not always representative of the prices at each generator due to congestion. This may have been a more significant problem in 2005.

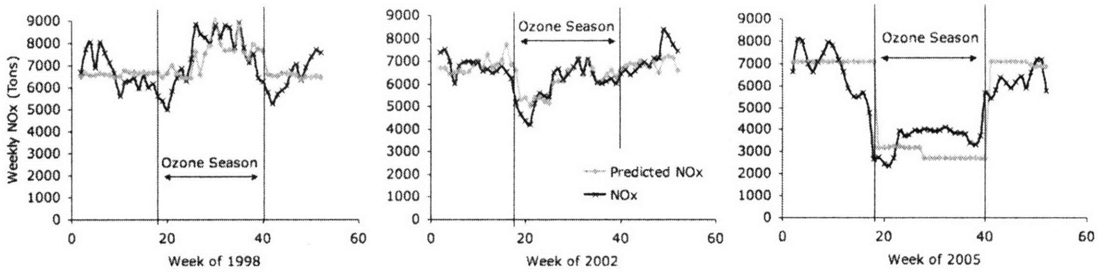

Figure C-6-5 Predicted and observed generation for coal units in the OTC region of the PJM and New England power systems in 1998, 2002, and 2005. The first stage of the predictions regressed generation on price with total electricity demand as an instrumental variable for price.

Table C-6-3 Predicted coefficients from regressions of generation on average electricity prices (with total electricity demand as an instrumental variable), heat input on predicted generation, and NO_X emissions on predicted heat input in 1998, 2002, and 2005 using weekly panel data for coal plants in the OTC region in the New England and PJM power systems.

1998

Coefficient	**Generation (gen)**	**Standard Error**	**Heat Input (HI)**	**Standard Error**	**NOx**	**Standard Error**
Price	761**	67	--	--	--	--
Early	-3028**	939	14485*	8084	-3.4	2.3
Late	-215	998	55161**	8025	-3.3	4.0
Constant	18273**	1282	-234143**	36150	24.9**	7.6
Predicted Gen	--	--	2.9**	0.53	--	--
Predicted HI	--	--	--	--	0.00019**	0.00005**
Predicted HI^2	--	--	--	--	6.5E-11**	2.9E-11**
Number of Observations		4440				
Number of Groups		88				

2002

Coefficient	**Generation (gen)**	**Standard Error**	**Heat Input (HI)**	**Standard Error**	**NOx**	**Standard Error**
Price	386**	44	--	--	--	--
Early	-2366**	612	12276**	4705	-12.3**	1.1
Late	-2187**	691	29808**	4133	-11.1**	1.4
Constant	14696**	1192	112319**	15052	-85**	39.2
Predicted Gen	--	--	5.12**	0.60	--	--
Predicted HI	--	--	--	--	0.00089**	0.00029
Predicted HI^2	--	--	--	--	-1.3E-9**	5.44E-10
Number of Observations		6869				
Number of Groups		137				

2005

Coefficient	**Generation (gen)**	**Standard Error**	**Heat Input (HI)**	**Standard Error**	**NOx**	**Standard Error**
Price	281**	35	--	--	--	--
Early	2123**	702	-3034	4026	-34.5**	1.44
Late	-2010**	754	22994**	3530	-28.7**	1.30
Constant	34280**	2065	345741**	18078	-7075	4444
Predicted Gen	--	--	-0.01	0.35	--	--
Predicted HI	--	--	--	--	0.041	0.025
Predicted HI^2	--	--	--	--	-5.8E-8	3.57E-08
Number of Observations		4840				
Number of Groups		104				

** indicates significance at the 5% level
* indicates significance at the 10% level

Appendix D: Matching Generator Data to PJM FTR Load Flow Cases

The EIA and EPA generating unit-level data were first matched to the FTR network model by substation, by the name of the substation. But, within the substations it is sometimes difficult to determine which generating unit should be assigned to which generating bus. Some substations have over ten generating unit buses, which are all located at the same voltage level on the network. But, the buses are not identical because buses within each substation connect to lines with different characteristics and to different buses in the remainder of the PJM network. Given this type of ambiguity, a simple method was used to match the generating units to buses within each substation.

If the FTR model included generation data for a unit, the units were first matched to those with same hourly generation in the EPA database for days during the same time of year and level of demand as represented by the FTR model. Then, for the remaining units, those with the largest capacities in the EIA data were matched to those with the highest generation in the FTR model. If all the units in a substation had zero or the same generation in the FTR model, we matched the units to the EPA units at random. Some further changes to the locations of generators were made on a case-by-case basis as described in Appendix E.

Appendix E: Scaling the PJM FTR Load Flow Cases to Approximate Peak Demand Hours

The monthly FTR base case for July that PJM posted in July 2006 represents an hour with about 77 GW of load in PJM and 35 GW of load in Classic PJM. This was about average for the ozone season of 2005. This case is referred to as the "average demand" case for the PowerWorld simulations. The set of fossil-fuel-fired generating units in Classic PJM for which we have emissions data generated about 19 GW on average during the 2005 ozone season and also generated about 19 GW in this FTR case. Cost curve data were imported into PowerWorld for this set of generators; these are the generators that were redispatched and those to which the NO_X price applied. These generating units were referred to as the "Classic PJM fossil units" and in each case their total generation was held constant, the NO_X prices causes the reallocation of this generation to units with lower NO_X rates.

In order to simulate NO_X reductions from redispatch for an hour with peak conditions, and because detailed data on nodal loads during peak conditions were not available through the FTR cases, the load data in the average demand case were scaled.[161] This was not a straightforward exercise because loads on network do not scale uniformly from average to peak demand hours; the electricity demand increases more in some areas than in others. The patterns of load impact congestion and the corresponding LMP patterns. In order to simulate a peak hour with similar characteristics to observed peak hours the observed generation data from the fossil-fired units in Classic PJM were used (from the EPA's CEMS data) and the loads were scaled incrementally until LMP patterns and congested lines were observed that were reasonably similar to those observed in Classic PJM in the 2005 ozone season.

---

[161] Hourly load data are only available at the zonal level: PJM, "Hourly Load Data", available at http://www.pjm.com/markets/jsp/loadhryr.jsp.

Loads of over 50 GW and generation of about 35 GW from the Classic PJM fossil units characterized peak demand hours in Classic PJM in the summer of 2005. The EPA's CEMS data indicated that the Classic PJM fossil units generated about 33 GW on August 4th, 2005 at 2 pm – one of the peak hours. We used these unit-level data for the initial generation of Classic PJM fossil units and assumed that generation remained constant from nuclear, hydro and other units not included in the EPA's CEMS data. For the latter units, we used the generation data from the July FTR case. Holding the generation of these units constant is a reasonable assumption for the nuclear plants, which typically generate near their capacities in all hours, but may be a simplification for the other units. This assumption likely only created slight changes in the results because of the small contribution of these other units: EIA data suggest that hydro, wind, and fossil units for which we do not have data contribute about 5 GW of about 67 GW capacity in Classic PJM (about 7%). Additionally, if these units could respond to higher NO_X prices, excluding them from our estimates of potential NO_X reductions makes our estimates conservative.

The zonal model indicated that network congestion created about 11 zones with constant LMPs on August 4th 2 pm. In order to try to recreate these LMP and congestion patterns with PowerWorld, the factors by which the generation from Classic PJM fossil units in each zone increased between the average demand case and August 4th 2 pm observed data were calculated. The nodal loads in zones with higher than average LMPs were scaled by factors slightly higher than the corresponding generation scaling factor (and vice versa for zones with lower LMPs). It was likely that zones on the high-LMP sides of congested lines were net-importers of power because congestion reflects the fact that the high-LMP area was importing as much power as possible from remote generators with lower costs. The load in Classic PJM was scaled to about 53 GW using this method and imported the new load data into PowerWorld.

The inaccuracies inherent in the method of scaling the nodal loads and those in our method of matching generator data to the buses in the FTR case (Appendix D) caused some problems for solving the power flow simulations. The optimal power flow simulations dispatch generators, according to their marginal costs, to meet all loads while minimizing system operating costs and meeting a set of inequality constraints. These inequality constraints include generator and transmission line capacities. In our cases, we do not allow PowerWorld to dispatch the nodal loads: the generating units are the only "controls" that can be changed to minimize cost while holding demand constant. It is sometimes not possible for the generating units to be dispatched to both fill load and meet all the inequality constraints. In this situation, PowerWorld prioritizes filling the load and it reports "unenforceable" line or generator constraints. The flows on lines with "unenforceable" constraints exceed their rated capacities. These overloaded lines may actually occur in reality and be acceptable: lines' capacities vary with external conditions like the weather and the system operator has the discretion to adjust the lines' ratings in real-time while we do not have enough information to do so. Unenforceable constraints in the simulations may also be an artifact created by the inaccuracies of the load and generator data. If, for example, a generator were incorrectly placed at a bus it might cause large flows over a line that was not intended to handle them.

Unenforceable constraints were observed in both the average and peak cases. The locations of some of the generators and the magnitudes of the scaled loads were altered in the peak case to mitigate these problems, but were not all of the unenforceable constraints could be removed. Depending on the case, between zero and about twenty unenforceable constraints were observed out of over 3100 lines (less than 1%). Because of these inaccuracies, the simulation results are reported in relative terms (e.g. the higher NO_X-price cases are compared to the \$2,000/ton NO_X price case) rather than as absolute results compared to observed data for the modeled hour (e.g. August 4th at 2 pm).

Given the matching and scaling inaccuracies, the simplified cost curves, and the fact that PowerWorld does not allow for simulation of ancillary services markets, it was not possible to fully recreate historical conditions. The correct magnitudes of demand, fossil generation, NO_X emissions, and congestion were reproduced and therefore the results are reasonably representative of a power system similar to Classic PJM. From the scaled load and generation data described above created two peak load flow cases that reproduced the general characteristics of historical conditions by adjusting the nodal load data in the scaled case incrementally after loading the data into PowerWorld. One of these cases approximated the patterns of LMP observed on August 4th at 2pm ("Matched LMPs") and the other better approximated the observed contingency constraints ("Constraints"). The Classic PJM fossil units generated about 34 GW and emitted about 35 tons of NO_X in both simulations; the total load was about 53 GW. For comparison, on August 4th at 2 pm the observed total load in Classic PJM was about 59 GW and Classic PJM fossil units generated about 35 GW and emitted about 38 tons of NO_X. On August 3rd at 2 pm, another peak demand hour in PJM, the total load was about 54 GW and the fossil units generated about 33 GW and emitted about 35 tons of NO_X.

On August 4th at 2pm PJM reported active contingency constraints for five lines. Two of these were reproduced in the Matched LMPs case and four in the Constraints case, while assuming NO_X prices of $2000/ton. In addition, in the Matched LMP case the basic patterns of LMPs were recreated. For example, the constraint that caused the most variation in LMP in the observed hour was the Cheswold-Kent line. The nodes on the low side of this constraint had LMPs about 9% below average and those on the high side had LMPs about 60% above average (the observed mean LMP in this hour was $287/MWh). In the Matched LMPs case the nodes on the high and low side of this constraint also had the largest differences in LMP: 51% below average on the low side and 274% above average on the high side (the mean LMP in this simulation was $129/MWh). Although the magnitudes of LMPs are different between the simulated and observed cases, the patterns are similar. The inaccuracies discussed above contribute to these differences. In particular, generator cost curves typically increase steeply at high capacity factors and our simplified cost curves do not account for this.

In the third peak case, decreasing the assumed generation from other units increased the share of total generation from the Classic PJM fossil units. As in the other two peak cases, the total load was bout 53 GW, but in this case the fossil units in Classic PJM generated about 37 GW and emitted 39 tons of NO_X. This case was intended to simulate a worst-case scenario. The simulations hold the total generation from the Classic PJM fossil units constant and the NO_X reductions in response to the NO_X price can only come from the reallocation of generation between the units. Requiring these units to generate more reduces the flexibility to reduce emissions through redispatch because it reduces the excess capacity of the units that can be redispatched.

ACKNOWLEDGEMENTS

I am indebted to the members of my committee, Denny Ellerman, Paul Joskow, David Marks, and Kenneth Oye who have been so very patient and helpful throughout this process. Denny and Paul were the inspiration behind my research question. They have provided invaluable technical and expositional advice throughout and have been extremely generous with their time. A student could not hope for better research advisers, as they have been role models for professional collaborations that extend far beyond the research environment. Dave's ability to anticipate exactly which pep talk I needed at every stage of this academic journey has never ceased to amaze me and it carried me through difficult times. Ken's enthusiasm for research at the science-policy interface has been an inspiration since I started the TPP program five years ago. I have always marveled at his ability to make sense of the complicated issues at this boundary.

I am also indebted to the useful comments and encouragement given in response to presentations of this work at EPA, NOAA, the PJM Interconnection, and at the Energy and Environmental Policy Workshops of MIT CEEPR. I would like to thank Mark Gilrain and Andrew Ott at PJM, Nathan Lang and Andrew Stevens at DC Energy, Minghai Liu and Seabron Adamson at CRA International, Inc. for their advice about the research and for assistance with optimal power flow simulations.

I cannot possibly offer enough thanks to Mary Gallagher, Jackie Donoghue, Joni Bubluski, Therese Henderson, Elizabeth Milnes, and Sidney Miller for their invaluable assistance in navigating MIT, ESD and LFEE. Thanks especially to Jackie for always saying hi with a smile and to Mary for encouragement and for help with all the details. Conversations and classes with the many incredible MIT faculty members – especially in those in TPP – have made my time here rewarding and challenging. I would also like to thank Larry McCray for his wisdom and for always asking all the most interesting research questions. I owe thanks to Steve Connors as well for first bringing my attention to the power of detailed historical data.

To my fellow TPP, TMP, and ESD students and my E40 cubicle mates who have come, gone, or are still here: thank you for inspiring me. MIT has blessed me with so many amazing friends. Julia, Emma, Fiona, Mike, Jenny, Tom, Michelle, Mary, Anup, Chris, Adam, Rogol, Alex, Travis, and the other Kate: thanks for going through this with me and for all the fun times and distractions. Thanks also to Jason, Ralph, Christine, Mark, Erica, Jim, Marcus, and Mort for going through this before me and for all your advice.

Finally, thank you to my wonderful family and especially to my parents, brothers and to John for being so supportive and loving and for always reminding me of the 'funner' and funnier sides of life. I can't wait to enjoy more of them with you.

I would like to dedicate this dissertation to my mom, for teaching me how to write.