Empirical Studies of New Markets in the US Electricity Industry

by

Elizabeth M. Bailey

B.A. in Mathematics, Colgate University, 1994

Submitted to the Department of Economics
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 1998

©1998 Elizabeth M. Bailey. All rights reserved.

The author hereby grants to MIT permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author...

Department of Economics
April 15 1998

Certified by...

Paul L. Joskow
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Certified by...

Richard L. Schmalensee
Gordon Y. Billiard Professor of Economics and Management
Thesis Supervisor

Accepted by...

Peter Temin
Chair, Departmental Committee on Graduate Students
Empirical Studies of New Markets in the US Electricity Industry

by

Elizabeth M. Bailey

Submitted to the Department of Economics
on April 15, 1998 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

ABSTRACT

This thesis presents empirical studies of two new markets in the United States electricity industry. Chapter one focuses on the wholesale electricity market in the western United States. Chapters two and three focus on the tradable permits program created by Title IV of the 1990 Clean Air Act Amendments.

Chapter one develops new techniques to assess the expanse of the geographic market under varying supply and demand conditions and applies these techniques to the current wholesale electricity market in the western United States. This paper finds that, by and large, the expanse of the geographic market extends across most of the western United States, but that conditions which create congestion along transmission lines, such as high hydroelectric flows in the Pacific Northwest, transmission line outages and deratings, and high demand for wholesale electricity, cause the expanse of the geographic market to narrow at certain times.

Chapter two explores whether regulatory activity at the state level complicates electric utilities' decision to trade sulfur dioxide emission allowances as allowed by Title IV of the 1990 Clean Air Act Amendments. This paper finds that public utility commission regulation has encouraged allowance trading activity in states with regulatory rulings, but that allowance trading activity has not been limited to states issuing regulations.

Chapter three presents the inter-temporal pattern of allowance prices that should be observed in the market for sulfur dioxide allowances in world of certainty with no transaction costs, and demonstrates that this pattern is roughly consistent with what is observed. The empirical analysis in this paper suggests that the forward market for emission rights has become reasonably efficient by early 1996.

Thesis Supervisors: Paul L. Joskow, Elizabeth and James Killian Professor of Economics
Richard L. Schmalensee, Gordon Y. Billiard Professor of Economics and Management
Acknowledgements

I have found the colleagues that I have been privileged to work with here at MIT one of the most enjoyable and rewarding aspects of the MIT economics department.

I am most grateful to have had two advisors who were very generous with their time and with their experience. It has been a pleasure and a privilege to have had the opportunity to work for, and later with, Paul Joskow and Dick Schmalensee. Paul has been a constant source of advice and encouragement during my four years at MIT. My sincerest thanks go to Paul for sparking my research interests and shaping my thinking as an economist. I continue to admire Paul’s careful and thoughtful approach to empirical research, and I would be pleased to find that my own research reflects his influence. Richard Schmalensee has shown an amazing ability to find clarity in my muddled thoughts. His willingness to help by providing timely, detailed, and generous comments has improved my thinking at every stage. I am also grateful to Nancy Rose for providing insights and encouragement which improved this thesis immensely. I also want extend my deepest thanks to Dr. Denny Ellerman, the executive director of the MIT Center for Energy and Environmental Policy Research, for providing me with numerous opportunities to present my research, the opportunity to interact with a wonderfully diverse and bright group of people, and for very generously providing me with financial support over most of the course of these four years.

I also feel very lucky to have made several dear friends here at MIT. I cannot begin to thank Paul Ellickson, Marianne Bitler, and Ezra Friedman for all their friendship and the many kindnesses that they have shown. I have learned from them, both about economics and about myself. For this reason, it is with great sadness that I watch these four years come to an end.

Most of all, I want to thank all of my family for encouraging me and being proud of me.
Contents

1. The Geographic Expanse of the Market for Wholesale Electricity 5
   1.1 Introduction. .................................................. 5
   1.2 Spatial Price Dispersion in the Western US Wholesale Electricity Market. . 10
   1.3 Formalizing Wholesale Price Relationships. .................................. 16
   1.4 The Data ..................................................................... 20
   1.5 Empirical Examination of the Expanse of the Market. ...................... 22
       1.5.1 The Effect of Regional Supply and Demand Conditions
             on Price Correlations ............................................... 23
       1.5.2 The Effect of Congestion on Arbitrage Constraints .................... 29
       1.5.3 How Frequently Do Arbitrage Constraints Bind. ....................... 37
   1.6 Conclusions and Implications. ................................................. 44

2. Allowance Trading Activity and State Regulatory Rulings: Evidence
   From the US Acid Rain Program 65
   2.1 Introduction. .......................................................... 65
   2.2 Allowance Trading Activity and PUC Regulations ............................ 68
   2.3 Empirical Specification. ................................................ 73
   2.4 Empirical Results ................................................................ 78
   2.5 Additional Evidence on the Effect of State PUC Regulation ............... 81
   2.6 Implications and Conclusions. .............................................. 84

3. Intertemporal Pricing of Sulfur Dioxide Allowances 100
   3.1 Introduction. .......................................................... 100
   3.2 Title IV and Inter-temporal Allowance Trading. .............................. 103
   3.3 Theoretical Framework. .................................................. 107
   3.4 Empirical Evidence. ...................................................... 111
   3.5 Concluding Observations. ................................................. 115
CHAPTER 1

The Geographic Expanse of the Market for Wholesale Electricity

1.1 INTRODUCTION

Spatial price relationships have been used widely in the literature as a means to infer the geographic expanse of the market, and more generally, market performance. A large number of goods have been the subject of previous studies of the geographic expanse of the market. For example, Horowitz (1981) examined the geographic expanse of the market for meat products sold for consumption in the eastern United States. Elzinga and Hogarty (1973) considered the geographic extent of the market for beer. Several markets including those for wholesale flour, wholesale unleaded gasoline, and residual fuel oil captured the attention of Stigler and Sherwin (1985) in their approach to estimating the geographic expanse of the market. Both Spiller and Huang (1986) and Spiller and Wood (1988) examined the geographic expanse of the market for wholesale gasoline in the northeastern United States. The geographic expanse of the market for agricultural products is especially well represented in the literature including studies of the market for California tomatoes, Canadian hogs, California and Florida celery, Nigerian food grain products, and rice in Bangladesh.¹

Although each study examines the geographic expanse of the market for a different good using slightly different empirical methods, all these empirical approaches share several common features. First, transportation costs as well as all other characteristics of the good are treated as fixed over the time period of the data. Second, changes in the direction of flow of trade are usually ignored in the analysis, so that the expanse of the market is considered to be the same whether the good flows from region A to region B or whether the good flows from region B to region A. Third, the expanse of the market is treated as a pair-wise relationship (region A and region B are treated independently of any third region C), that is to say, any network characteristics of the good are ignored in the analysis. Fourth, with the exception of Spiller-Huang and Spiller-Wood, the empirical methodologies used do not provide results which shed light on how frequently the regions are in the same geographic market and how frequently the regions are not in the same geographic market. Moreover, none of the methodologies distinguish between the two very different reasons that a region could be “out of the market”: economical reasons could keep a region “out of the market” (i.e. autarky) as well as inefficiency reasons (i.e. congestion). Taken together, these features of the methodologies provide only a very static description of the geographic expanse of the market.

The spatial dispersion of prices, however, is rarely static. First, transportation costs, as well as other characteristics of the good, change over time, and, as a result, the geographic expanse of the market changes over time as well. Second, the direction of flow of trade may change over time in response to changing marginal costs of production. As a result, the geographic expanse of the market may look very different when the good flows in a north to south direction than when the good flows in a south to north direction. Third, if trade occurs on a network, constraints to
trade along one path will have implications for the expanse of the market elsewhere in the network. For these reasons, spatial price dispersion changes over time, and as a result, it is unlikely that the geographic expanse of the market for a particular good can be characterized once and for all as a single specific size. Rather, the geographic expanse of the market for a particular good is more appropriately though of as "wide" under some conditions, "narrow" under other conditions because trade between regions is not economical, and finally, in still other conditions, "narrow" because congestion prevents economical trade from occurring. In sum, the geographic expanse of a market for a particular good is dynamic.

The wholesale\(^2\) electricity market in the western United States is particularly well suited to a study of the dynamics of spatial price dispersion and, more particularly, how the geographic expanse of the market changes over time. First, electricity is a non-storable good. Second, the market must equilibrate instantaneously. Third, wholesale electricity trade occurs on large network of high voltage transmission lines that link the many regions together. As a result, shocks to supply and demand in one region are likely to have noticeable effects on wholesale prices in other regions. Finally, wholesale energy trade between vertically integrated utilities in the western U.S. has grown over time and has become increasingly more sophisticated. Initially, wholesale energy trade typically involved transactions between two physically interconnected utilities that utilized their own transmission capacity to consummate bilateral transactions. Until 1992, utilities were not obligated to supply "unbundled" transmission or wheeling service to others to support wholesale transactions, although some utilities provided such service

\(^2\) Wholesale electricity transactions are purchases and sales of electricity by electric utilities to each other. Retail electricity transactions are sales to end-use residential, commercial, and industrial customers.
When unbundled transmission service was not available, transactions involving trade over larger geographic areas often involved a series of “buy/sell” transactions involving three or more utilities. Over the last decade, unbundled transmission service, especially for short term transactions (daily and hourly) has become increasingly available in the western US as a result of voluntary multilateral agreements, requirements imposed on vertically integrated utilities by the Federal Energy Regulatory Commission (FERC) to file open access transmission tariffs as a condition for approval of merger applications and application to sell wholesale electricity at market based rates, and most recently through the obligations created by FERC Order 888 (effective January 1 1997) which require all vertically integrated utilities to file open access transmission tariffs that meet certain criteria specified by FERC under authority provided by the Energy Policy Act of 1992. The availability of unbundled transmission service has increased opportunities for bilateral energy trade among utilities within the same region, but which are not necessarily directly interconnected. As a result, a study of the dynamics of spatial price dispersion can shed light on how well integrated the wholesale electricity market in the western US has become as a result of increased opportunities for bilateral energy trade.

A study of the dynamics of spatial price dispersion in the current wholesale electricity market can also go a long way toward informing the discussion of pricing behavior and performance in a restructured electricity industry. The United States electricity industry is presently being restructured in order to promote competition in the supply of generation services at wholesale
and retail. An important issue in restructuring the electricity industry is the significance of seller market power associated with the current ownership patterns of generation capacity. One important input into a complete assessment of imperfect competition in a restructured electricity industry is the geographic expanse of the market for generation services. Because market conditions in the electricity industry are likely to change significantly in the next few years as the structure of the electricity sector changes dramatically, the present analysis can provide a useful benchmark against which to compare post-restructuring wholesale price relationships.

In this paper, three techniques are developed to assess how the geographic expanse of the market for wholesale electricity changes in response to shocks to supply and demand. The first technique identifies the particular supply and demand conditions that give rise to a “narrower” market. This paper finds that high hydroelectric flows in the Pacific Northwest, transmission line outages and deratings, and high demand for electricity are conditions likely to give rise to congestion along transmission lines. The second technique assesses whether transmission line congestion causes arbitrage constraints to become non-binding. This paper finds that

3 On January 1st 1998, the three investor owned utilities in California will be restructured and retail customers will be permitted, for the first time, to buy power directly in the wholesale market or from competing brokers and marketers. Although the California restructuring project is by far the largest project to date, several other states have already instituted pilot programs which offer retail customers a choice of electricity suppliers. On August 1 1997, Rhode Island began offering large retail customers a choice of electricity suppliers. New Hampshire and Massachusetts have instituted pilot programs for smaller retail customers. The New Hampshire pilot program has been ongoing for about 12,000 customers since June 1996. Under the terms of the program, customers are free to purchase electricity from one of about 30 competitive power suppliers. The Massachusetts pilot program, “Choice: New England”, began in July 1997. The Massachusetts program offers 4,727 residential and small business customers a choice of electricity suppliers.

4 Wolfram (1996) finds some empirical evidence that generators exercise market power in the British electricity spot market. Using simulated data, Borenstein and Bushnell (1997) find evidence that generators in the soon to be restructured California electricity industry will be able to raise price above competitive levels. Joskow and Schmalensee (1983) find that the intensity of competition is likely to depend on the size of the relevant geographic market as determined by the distance over which electricity can be economically transmitted.
transmission line congestion does cause arbitrage constraints to become non-binding. The third technique assesses how often arbitrage constraints do not bind, providing insight into how frequently the geographic expanse of the market narrows. This paper finds that between June 1995 and December 1996, arbitrage constraints bind prices in eighty percent of the daily observations, transmission congestion causes price separation in nineteen percent of the daily observations, and autarky prevails in the remaining one percent of the daily observations.

The remainder of the paper proceeds as follows: Section 2 introduces the market for pre-scheduled wholesale electricity in the western United States and illustrates the dynamic nature of the wholesale electricity market. Section 3 formalizes how the market for pre-scheduled wholesale electricity equilibrates. Section 4 describes the data. Section 5 presents the empirical investigation of how regional supply and demand conditions affect the arbitrage constraints between sub-regions of the western United States and cause the geographic market to widen or narrow. Section 6 provides conclusions.

1.2 SPATIAL PRICE DISPERSION IN THE WESTERN US WHOLESALE ELECTRICITY MARKET

The Western System Coordinating Council (WSCC), the electricity market that this paper focuses on, covers all of the contiguous US states west of the Rocky Mountains, two provinces in Western Canada (British Columbia and Alberta) and portions of Northern Mexico (See Figure 1). The electric power companies located in this large region operate as part of a single synchronized AC network, known as the Western Interconnection, whose physical operation and reliability is coordinated by the WSCC, a voluntary regional reliability council. It should not be of concern
that the effect of electricity systems in locations east of the Rocky Mountains has essentially been ignored. It is not likely that supplies from the east and southeast could have a significant effect on prices in the WSCC since only small DC interconnections connect the WSCC with the Eastern Interconnection and the Texas Interconnection, the two other synchronized AC systems operating in the United States.

The WSCC region can be sub-divided into five smaller regions⁵ - Northwest, Northern California, Southern California, Inland Southwest, and Central Rockies - which reflect the distinct geographic and climatic conditions in the WSCC.⁶ By and large, the bulk of the electricity supplied to retail customers in the WSCC is supplied by vertically integrated electric utilities that serve retail loads in exclusive geographic areas using primarily own-generation, but because high voltage transmission link these five sub-regions together, utilities in the five regions are able to buy and sell electricity at wholesale under bilateral contracts that cover exchanges of energy for periods as short as an hour and as long as ten years. These wholesale electricity transactions allow vertically integrated utilities to trade on the margin with one

---


⁶Hydroelectric facilities constitute a significant part of the resources available to serve demand especially in the Pacific Northwest and Northern California. Electricity demand in the Pacific Northwest peaks in the chilly winter months, while Northern California electricity demand peaks during the warm summer months. Generating capacity in Southern California is predominantly nuclear and natural gas. Demand for electricity in Southern California peaks in the hot summer months. Coal-fired and nuclear are the largest contributors to generating capacity in the Inland Southwest. Like California, demand for electricity in the Inland Southwest peaks in the hot summer months. Generating capacity in the Central Rockies is predominantly coal-fired. Electricity demand peaks in the summer months as well. These five sub-regions correspond to the five major supply and demand areas in the WSCC. Electricity demand in these five regions is not perfectly correlated because of variations in weather conditions (temperature as well as precipitation) and time zone changes (the WSCC covers states located in the Pacific Time Zone as well as states located in the Mountain Time Zone).
another to reduce the utility's cost of meeting its own retail electricity demand. The majority of transactions for wholesale electricity in the WSCC are daily, pre-scheduled transactions. The Table 1 presents data from the WSCC on 1995\(^8\) annual aggregate quantities of short-term imported (as opposed to self-generated) and exported electricity.

High voltage transmission lines link the five sub-regions of the WSCC together. Figure 2 illustrates the location and relative size, as measured by the maximum MW capacity of the transmission line, of the major high voltage transmission lines in the WSCC. In the WSCC, there are competing transmission paths for most routes. For example, under most conditions, utilities can transport wholesale electricity from the Northwest into Southern California either by the DC Pacific Intertie, which directly connects the Northwest with Southern California, or by using a combination of the AC Pacific Intertie, which connects the Northwest with Northern California, plus the lines at Midway which connect Northern California to Southern California. Most transmission lines, including the three major lines (the AC Pacific Intertie, the DC Pacific Intertie, and the AC lines which connect Southern California with the Inland Southwest), are owned by several utilities each which has ownership shares of the line’s transmission capacity. These entities own the transmission rights to the transmission line and have the right to use the capacity themselves or to sell the transmission rights to other utilities or third parties.

\(^7\) Pre-scheduled energy transactions are energy transactions that are scheduled a day-ahead of actual physical production.

\(^8\) 1996 data are not available.
The capacity of a transmission line is limited. The maximum allowable rating of the line defines the maximum megawatts of electricity that the transmission line can transport and therefore the maximum amount of wholesale electricity trade that can occur between the sub-regions connected by that line. Figure 3a and 3b illustrate graphically the daily ratings of the southbound AC Pacific Intertie and of the southbound DC Pacific Intertie between June 23 1995 and May 9 1997. It is not uncommon for a transmission line’s rating to be lowered below its maximum allowable rating, or “de-rated”, on a one day basis either for planned or unplanned maintenance or because of an unexpected line outage. Moreover, it is not uncommon for the rating of a transmission line to remain below its maximum allowable rating for extended periods of time as long-term maintenance is performed or for reliability concerns. Figure 3a and 3b demonstrate that the rating of the AC line and the rating of the DC line are constantly in flux over this period. When the transmission line constraint binds, that is the demand for wholesale electricity trade exceeds the rating of the transmission line, the transmission line is said to be “congested”.

The price for transmission service along a particular route is set competitively and varies as supply and demand conditions change throughout the WSCC. The price of transmission service depends positively on demand for transmission service: Variables that increase demand for wholesale electricity trade also increase the demand for transmission service. FERC regulates the price of transmission service using a price cap. The maximum price a utility can charge for transmission service is defined by the average embedded cost of the utility’s network.\footnote{Averaging across all utilities in the WSCC, the regulated price cap for hourly, non-firm, point to point transmission service is 3.69Mills/Kwh with a standard deviation of 1.6Mills/Kwh. The lowest price cap set by FERC is 1.33Mills/Kwh (Idaho Power Company). The highest price cap set by FERC is 7.1Mills/Kwh (San Diego Gas and Electric).}
result, when there is congestion the competitive price for transmission service may not always be able to rise to clear the market.

Figure 4a (4b) present the daily peak (off-peak)
10 period pre-scheduled wholesale prices in the five sub-regions of the WSCC from June 23 1995 to December 31 1996. Both Figure 4a and Figure 4b illustrate that wholesale electricity prices in the five sub-regions of the WSCC move more or less together, consistent with the hypothesis that the wholesale electricity market is a single market westward from the Rocky Mountains. Although the wholesale prices appear to move in tandem, Figure 4a and 4b also indicate that there is quite a bit of time of day (peak versus off-peak) and seasonal fluctuation in prices. These time of day and seasonal fluctuations in prices reflect changes in marginal costs due to available generation resources as well as changes in the quantity of electricity demanded by retail consumers. Figure 4a and 4b also illustrate an important attribute of electricity: electricity does not always flow in the same direction (e.g. from region A to region B). For example, in the WSCC during daytime hours in the Spring, wholesale energy typically flows from the Northwest into Northern and Southern California because of the abundant inexpensive hydroelectric resources available at that time in the Northwest (P_{NW} > P_{NCA} and P_{NW} > P_{SCA}). In the late Fall and early winter though, utilities in California are more likely to be exporting electricity to the Northwest, especially at night when it is cold in the Northwest, in order to conserve scarce hydroelectric resources in the Northwest for use during peak hours (P_{NW} < P_{NCA} and P_{NW} < P_{SCA}). Figure 5 presents examples of the geographic

---

10 Peak: daytime weekday hours (6:00am-9:09pm); Off-peak: nighttime hours (first six and last two hours of the day) and all hours on weekends and national holidays. Price for Saturday and Sunday are scheduled by control area operators as if they were the same day.
dispersion of wholesale electricity prices on typical days when electricity flows in different directions.

Although Figure 4a and Figure 4b illustrate that wholesale electricity prices in the five sub-regions of the WSCC move more or less together suggesting that the geographic market is quite wide much of the time, it is clear from Figure 4a and 4b that there are time when prices do not move together suggesting that under some conditions the geographic market may be much “narrower” than the entire expanse of the WSCC. For example, prices in the Northwest and the Inland Southwest seem to separate on September 7 1995.\(^\text{11}\) Figure 6 highlights the wholesale electricity prices between September 6-8 1995 from Figure 4a. September 7 1995 was marked by unusually hot temperatures in the Inland Southwest. The high temperatures increased retail cooling loads and, as a result, increased demand for wholesale electricity imports in the Inland Southwest from the Northwest. The transmission lines into the Inland Southwest from the Northwest were filled to capacity and unable to accommodate the increased demand for wholesale electricity. As a result, utilities in the Inland Southwest were forced to rely on more expensive own generation resources causing the price for wholesale electricity to rise in the

\(^{11}\) In general, congestion causes the price difference between two regions to be greater than the regulated transmission price cap by only a few Mills. On occasion though, congestion can cause regional price differences to be well in excess of the regulated transmission price cap. For example, on August 5 1997, a day of record high retail electricity demand due to hot weather, a small plane clipped two 500 KV lines in Southern California Edison’s system. The downed transmission lines created a cascade of unit outages in and around Southern California, and in addition, made it difficult, if not impossible, to import power into Southern California. As a result, brown outs were reported all around the Los Angeles area. Prices for pre-scheduled power for August 7 1997 (scheduled on August 6 1997) rose by gargantuan amounts in all regions as Southern California and the Inland Southwest tried to cover loads by purchasing wholesale power from the Northwest and the Central Rockies. Prices in Southern California and the Inland Southwest peaked as high as 100Mills/Kwh (up 55Mills/Kwh from the day before the disturbance) and 160Mills/Kwh (up 113Mills/Kwh from the day before the disturbance) respectively. Prices in the Northwest and Central Rockies were reported as high as 37Mills/Kwh (up 15.5Mills/Kwh from the day before the disturbance) and 60Mills/Kwh (up 31Mills/Kwh from the day before the disturbance).
Inland Southwest. The demand shock had no effect on wholesale electricity prices in the Northwest because the transmission constraint prevented more energy from flowing from the Northwest into the Inland Southwest. As a result, wholesale electricity prices in the Northwest remain virtually unchanged. Because it was physically impossible for additional trade to occur between the Northwest and Inland Southwest even though it appears economical to engage in wholesale transactions, the Northwest and the Inland Southwest are properly conceptualized as a separate geographic market on September 7th. Just as quickly as pre-scheduled wholesale prices in the Inland Southwest rose sharply in response to the heat wave, pre-scheduled wholesale prices across the WSCC fell back to levels observed prior to the demand shock as the heat wave dissipated. The next section formalizes precisely how the pre-scheduled market for wholesale electricity equilibrates and in what sense the sub-regions of the WSCC are conceptualized as a “wide” or “narrow” geographic market.

1.3 FORMALIZING WHOLESALE PRICE RELATIONSHIPS

Although the nature of energy flows are complex, in general, electricity flows from low “autarky” price regions to high “autarky” price regions until the gains from trade are exhausted unless transmission constraints or local operating requirements prevent further economical trading opportunities. Assuming a competitive market for generation services and transmission

---

12 The current wholesale electricity market in the WSCC is relatively competitive. First, there are a large number of generation suppliers at all locations around the network. Under normal, uncongested periods, wholesale electricity customers can turn to generation suppliers dispersed over a wide geographic area purchasing or selling electricity either within their control area, in directly connected control areas, or in more remote control areas. When congestion constrains the geographic expanse of the market, wholesale electricity customers are still able to turn to a large number of generation suppliers located within their control areas, including self generation. Second, the ability of sellers to raise price above competitive levels is constrained by the fact that most of the buyers of wholesale electricity are vertically integrated utilities which own their own generating capacity, and therefore have the opportunity to generate electricity themselves. Finally, sellers have little incentive to restrict output in order to
service and no congestion on the transmission line connecting region A and region B, Figure 7a demonstrates graphically how the market for wholesale electricity equilibrates. Suppose the demand for retail electricity in region A and region B are $Q_A$ and $Q_B$ respectively. Without wholesale electricity trade, the marginal cost to generate $Q_A$ units of electricity to fill retail demand in region A is $MC_{A, autarky}$ and the marginal cost to generate $Q_B$ units of electricity to fill retail demand in region B is $MC_{B, autarky}$. Suppose it is the case that region A’s “autarkic” marginal cost is greater than region B’s “autarkic” marginal cost ($MC_{A, autarky} > MC_{B, autarky}$). If the difference between the “autarkic” wholesale electricity prices is less than the competitive price of transmission service from region B to region A, wholesale electricity trade is not economical and utilities fulfill their own retail electricity demand using only their own generation resources. Suppose, though, the difference in autarkic prices between the two regions is greater than the competitive price for transmission service to transport electricity from region B to region A. In this case, engaging in wholesale trade reduces the cost of meeting retail electricity demand in each region. When there is no congestion, utilities in each region engage in trade (of the amount $X$) until the gains from trade are exhausted region (region A gains area $\Pi_A$ and region B gains area $\Pi_B$). The difference in the prevailing prices for wholesale electricity is the competitive price

---

raise price because they can only make a profit by expanding output beyond what is required to supply their retail customers at regulated prices.

13The market for transmission rights in the WSCC is quite competitive. Most transmission lines, including the two major lines are owned by several utilities each which has ownership shares of the line's transmission capacity and can either use the line themselves or sell the rights to use the line to other utilities or third parties. In addition, there are competing transmission paths for most routes.
for transmission service: $P_A - P_B = T_{BA}$.

In this case, the geographic expanse of the market for wholesale electricity encompasses both region A and region B.

Suppose the transmission line between region A and region B is not fully utilized at the equilibrium $P_A$ and $P_B$. If there is an increase in demand for retail electricity in region A, the autarkic marginal cost in region A rises. Wholesale electricity imports from region B to region A increase to give a new equilibrium, $P'_A$, $P'_B$, and $X'$ (see Figure 7b). If the marginal cost of transmission service is constant, then equilibrium wholesale prices in region A and region B increase by the same amount, $P'_A - P'_B = T_{BA}$. The price in region B responds to the demand shock in region A and the geographic expanse of the market for wholesale electricity continues to encompass both region A and region B.

Now suppose at the equilibrium $P_A$, $P_B$, and $X$, the transmission line between region A and region B is just fully utilized. In this case, a demand shock in region A which increases prices in region

---

14 This case corresponds to the framework which Schweppes et al (1988) have in mind in their proposition that the short run marginal costs of generating electricity at each node on the network are efficient locational spot prices, and the difference in the locational spot prices at any two nodes on the network is, subject to capacity constraints, the marginal cost of transmission between those two nodes.

15 In the WSCC, a constant marginal cost of transmission service is a reasonable assumption because reliability, rather than thermal constraints, limit transfers along the major transmission paths in the WSCC. For this reason, the marginal cost of transmission service is relatively flat up to the rated limit.

16 The analysis of the geographic expanse of the market becomes more complicated if the marginal cost of transmission service is upward sloping. If the marginal cost of transmission service is upward sloping, some of the price increase in region A will be due to an increase in the competitive price for transmission service and the rest will be due to an increase in the price for wholesale electricity in region B: $P''_A - P''_B = T_{BA}$. The difference in the wholesale prices between the two regions will increase to reflect the increase in transmission prices ($T'_{BA} > T_{BA}$). Prices between the two regions will appear to be less closely related even though the geographic expanse of the market for wholesale electricity continues to encompasses both region A and region B, in the sense that a demand shock in region A causes a price response in region B. This paper controls for the possibility that transaction costs may be different under different conditions by assessing price relationships separately under various sets of supply and demand conditions.
A will have no effect on wholesale electricity prices in region B because the transmission constraint prevents more energy from flowing from region B into region A (Figure 7c).\(^{17}\) If the price for transmission service, \(T_{BA}\), is unregulated, \(T_{BA}\) will rise in order to allocate the scarce transmission capacity and arbitrage constraints will continue to bind. The increase in the price in region A is reflected entirely in the competitive price of transmission service. If the price for transmission service is instead regulated by a price cap, as it is in the WSCC, the increase in the price in region A may not be able to be fully reflected in the price of transmission service. The wholesale electricity price in region B continues to be unaffected by the demand shock in region A, but now the price of transmission service is able to rise only as high as the regulated price cap.

In the case when the price of transmission service is regulated by a price cap, arbitrage constraints will not bind: the price in region A will exceed the price in region B plus the price of transmission service from region B to region A. In either the regulated or unregulated case, when there is congestion each region is properly conceptualized as a separate geographic market because the wholesale electricity price in region B is unaffected by the demand shock in region A.

This suggests that when the price of transmission service is regulated by a price cap, as it is in the wholesale electricity market in the WSCC, that there are two ways to identify the effect of congestion on the geographic expanse of the market. First, demand shocks in region A which increases prices in region A will have no effect on wholesale electricity prices in region B when the transmission line from region B to region A is congested. Second, arbitrage constraints will

\(^{17}\) The price in region B cannot rise because generation services in region B are competitive.
not bind when congestion is present. If the price of transmission service is unregulated, only the first technique can be used to identify the effect of congestion on the geographic expanse of the market.

1.4 THE DATA

Daily transaction price data for pre-scheduled, non-firm\(^{18}\), wholesale energy traded during “peak” and “off-peak” periods in the five sub-regions of the WSCC for the period covering June 1995 to December 1996 form the basis of the empirical analysis. These data have been provided by Economic Insights, Inc., a firm which continuously surveys electricity market transactions in the WSCC and makes the information available (for a fee) to utilities and intermediaries engaged in energy trade in the WSCC. During the course of the day, the Economic Insights staff contacts by phone a subset of the over fifty different utilities, marketers, and power brokers that have agreed to exchange information with the staff. In the average week each source is contacted two or three times. To assure correct information has been relayed, the staff seeks confirmation from at least two sources before including the data in the survey. Specific transactions involving names of the parties involved, quantities, and prices are confidential. Prices reported by Economic Insights are transacted prices (as opposed to bid or offer prices), including the price of transmission when appropriate, as reported by the participants contacted that day for non-firm, pre-scheduled power.

\(^{18}\) Firm service refers to non-interruptible service. Non-Firm service refers to service which may be interrupted.
Prices for each sub-region\textsuperscript{19} are broken into two categories: peak energy and off-peak energy. For each category two prices are reported: the high and the low.

The Economic Insights data have been transformed for the purposes of this paper. In this paper, a single day of price information consists of two price observations (peak and off-peak) for each of the five sub-regions. Whereas the Economic Insights data report a high and a low price for each sub-region in each period, this paper simply takes an average of the high and the low price to arrive at an average transaction price for each sub-region in each period. For example, if Economic Insights reports that in the peak period in the Northwest the low price was 10 Mills/Kwh and the high price was 13 Mills/Kwh, this paper would report the transaction price in the Northwest in the peak period to be 11.5 Mills/Kwh.

The data used for this paper cover the period June 23\textsuperscript{rd} 1995 to December 31\textsuperscript{st} 1996. Over this period, there are 474 days with price observations for the five sub-regions in the WSCC. Each non-holiday weekday has two observations for each sub-region: one peak and one off-peak. Each weekend and holiday weekday has two observations for each sub-region: two off-peak. In total, each sub-region has 948 price observations: 379 peak period observations and 569 off-peak period observations.

\textsuperscript{19} The wholesale prices reported by Economic Insights Inc. for each sub-region are in fact an average price throughout that entire sub-region. In practice the price of wholesale electricity varies even within each of the five sub-regions.
Supplemental data on daily supply and demand conditions thorough the WSCC are used to augment the price data. Information is available on daily maximum allowable capacities of transmission lines, unexpected transmission line outages, reports of transmission line congestion and congestion on the network, generating plant outages (unplanned, planned repairs, and economic outages), natural gas prices, contract coal prices, water flows at key locations in the Northwest, realized daily peak and off-peak demand for retail electricity. In addition, daily readings of maximum and minimum temperature, precipitation, and snow depth as well as 30 year norms for temperature, precipitation and snow depth are available for major cities across the WSCC.

For the bulk of the day ahead energy transactions that are analyzed there are no observable prices for transmission service that can be matched directly with the prices paid for energy. The empirical techniques developed below estimate the implicit price for transmission service. The regulated price cap for transmission service is used to check the accuracy of the estimation. Estimated prices for transmission service between two regions should not exceed the value of the regulated price cap if arbitrage constraints bind the two regions together. When there is congestion on the transmission line connecting two regions, the estimated difference in regional wholesale prices in these two regions should exceed the regulated price cap for transmission service.

1.5 EMPIRICAL EXAMINATION OF THE EXPANSE OF THE MARKET
1.5.1 The Effect of Regional Supply and Demand Conditions On Price Correlations

If two or more regions are in the same geographic market, prices in the two regions should move together quite closely: Shocks to supply or demand in one region should be transmitted to all other integrated regions. If either autarky or congestion prevails in two or more regions, prices in the two regions will be less tightly related. In general, shocks to supply or demand which increase congestion on the network narrow the geographic expanse of the market. The effect that these shocks have on the expanse of the geographic market should be reflected in the magnitude of the price correlations.

A methodology is developed in the spirit of the Stigler-Sherwin (1985) approach to explore how prevailing supply and demand conditions influence the expanse of the geographic market. Price correlations are generated from several categories of price observations and are regressed on the variables \( (X_k, k=1,2,...,K) \) which were used to divide the price data into the categories plus a set of dummy variables \( (PAIR_n, n=1,2,...,N) \) used to control for the particular market pair. The dependent variable is the price correlation for market pair \( n \) (e.g. Northwest and Northern California) for a particular set of price observations taken during times with factors \( \{X_1=x_1, X_2=x_2, ..., X_K=x_K\} \) present. The equation to be estimated takes the form:

\[
\rho_i = \sum_{n=1}^{N} \alpha_n PAIR_{i,n} + \sum_{k=1}^{K} \beta_k X_{i,k} + \epsilon_i
\]

where \( \rho_i \) is the price correlation for observation \( i \), \( PAIR_{i,n} = 1 \) if the correlation is from market pair \( n \) and 0 otherwise and \( X_{i,k} = 1 \) if the price correlation is generated from the set of price data
observed in a period characterized by $X_k$ being present and 0 otherwise. The coefficient on PAIR should be positive and larger for price correlations from regions which are closer together relative to regions which are further apart since regions closer together are likely to have lower transaction costs and therefore are more likely to be bound by arbitrage constraints. Price correlations generated from prices observed in times with factors which increase congestion present should be lower than price correlations generated from prices observed in times without congestion factors present. Consequently, the coefficient on $X_k$ is expected to be negative if $X_k$ increases congestion.

The standard Stigler-Sherwin test of market expanse needs to be carefully interpreted because price correlations fail to account for common supply and demand shocks that may lead prices to move together even when there are no arbitrage opportunities between the regions where the prices are observed.\(^{20}\) The approach developed in this Section is not fraught with that difficulty. Factors that may lead prices in different regions to move together, for example, common inputs to generating units such as coal and natural gas, are common to all categories. The technique developed in this Section relies on the change in the level of the price correlation not the level of the price correlation to interpret how different supply and demand conditions affect the expanse of the geographic market.\(^{21}\)

---

\(^{20}\) For example, two or more regions may exhibit high price correlations even when transmission lines connecting the regions are severed (clearly a case of market separation) if generating units across the regions rely on common inputs factors such as coal and natural gas. Changes in the common input prices would affect all regions similarly yielding a false positive result of market integration.

\(^{21}\) The typical fix to correct for the effect of common supply or demand factors is to run price correlations for first or longer price differences. This technique was not used due to the nature of the price data: peak and off-peak observations are not necessarily consecutive. Price data from a non-holiday weekday consist of one peak and one
To apply this technique, I divide the set of 948 price observations into mutually exclusive and exhaustive categories based on several variables.\textsuperscript{22} The variables are: Season (SPRING, SUMMER, FALL, WINTER), time of day (PEAK), reports that transmission lines from the Northwest in Northern California (NWNCAF) or from the Northwest in Southern California (NWSCAF) are full, a low transmission line rating on the Pacific Intertie (DERATING)\textsuperscript{23}, and a high hydroelectric flow at the Chief Joseph Dam in the Pacific Northwest (CJDAMH)\textsuperscript{24}. For example, one category is the set of prices such that it was Spring (SPRING=1), transmission lines from the Northwest into Northern California were not full (NWNCAF=0) but lines into Southern California were full (NWSCAF=1), the rating of the Pacific Intertie was low (DERATING=1), and hydroelectric flows along the Columbia river were high (CJDAMH=1). These variables generate a total of $128^{25}$ mutually exclusive and exhaustive categories.

Each of these variables can be associated with transmission line congestion and therefore can have an effect on the expanse of the geographic market for wholesale electricity in the WSCC.

\textsuperscript{22} Using too many variables to partition the data is problematic. In the limit, each category would have only one observation in it, too few to compute a price correlation.

\textsuperscript{23} During the period that the data covers, June 23\textsuperscript{rd} 1995 - December 31\textsuperscript{st} 1996, the maximum allowable rating of the combined Pacific Intertie (AC line plus DC line) was 7900 megawatts (MW). This rating fluctuated over the time period from 7900MW to 2000MW. A low rating on the Pacific Intertie is designated to be a rating below 6000MW. The results in Table 3 are invariant to small changes around this cut-off.

\textsuperscript{24} During the period that the data covers, June 23\textsuperscript{rd} 1995 - December 31\textsuperscript{st} 1996, hydroelectric flows along the Columbia river fluctuated between 43.7kcf and 197.1kcf. A high river flow is designated to be a flow above 150.0kcf. The results in Table 3 are invariant to small changes around this cut-off.

\textsuperscript{25} $128=4*2*2*2*2$ (4 seasons times 2 states for each of the other 5 categories).
First, since the Spring season is typically a time of high demand by Southern California and the Inland Southwest utilities for the cheap hydroelectric energy generated in the Pacific Northwest, demand for transmission capacity in the North to South direction is also high. As a result congestion tends to arise in a North to South direction during the Spring months. Second, peak periods are generally a time of high wholesale electricity demand and therefore high demand for transmission. As a result, congestion tends to arise in peak periods. Third, explicit reports that transmission lines are full suggests that transmission lines are likely to be congested. Fourth, when transmission line ratings are below their standard rated capacity, the line can carry less electricity between geographic regions. Consequently, transmission line congestion is more likely to arise. Fifth, when hydroelectric flows along the Columbia River in the Northwest are high, transmission lines from the Northwest region sometimes become congested in a southbound direction because the limited capacity of the transmission lines cannot accommodate all of the demand for the cheap hydroelectric power located in the Northwest.26

Price correlations are generated for each of the 128 categories of price observations. Eliminating the diagonal price correlations (those which equal one invariantly, i.e. the correlation between the Northwest and the Northwest), each category generates 10 price correlations. Price correlations generated from observations in categories with less than 2 observations -- 88 categories, of which 86 contained zero observations -- were eliminated. A total of 40027 price correlations remain.

26 The correlation between "spring" and "high hydroelectric flows" is 0.26.

27 128*10-88*10=400
The equation estimated by Ordinary Least Squares is:

\[ \rho_i = \sum_{n=1}^{10} \alpha_{i,n} PAIR_{i,n} + \beta_1 PEAK_i + \beta_2 SPRING_i + \beta_3 SUMMER_i + \beta_4 WINTER_i + \beta_5 NWNCAF_i \\
+ \beta_6 NWSCAF_i + \beta_7 DERATING_i + \beta_8 CJDMH_i + \epsilon_i \]

where \( \rho_i \) is the price correlation for observation i, \( PAIR_{i,n} = 1 \) if the price correlation is for market pair n, PEAK=1 if the price correlation is generated from peak period price observations, SPRING=1 if the price correlation is generated from spring price observations, SUMMER=1 if the price correlation is generated from summer price observations, WINTER=1 if the price correlation is generated from winter price observations, NWNCAF=1 the price correlation is generated from price observations taken during a time when the line connecting the Northwest to Northern California was reported full, NWSCAF=1 if the price correlation is generated from price observations taken during a time when the line connecting the Northwest to Southern California was reported full, DERATING=1 if the price correlation comes from price observations taken during a period when the cumulative rating of the southbound Pacific Intertie was low, and CJDMH=1 if the price correlation is generated from price observations taken from a period when the flows at the Chief Joseph dam were high.

Table 2 presents the parameter estimates from the regression. Because heteroscedasticity is introduced when the correlation coefficients are formed from different numbers of price observations, White corrected standard errors are reported. The price correlations for the ten market pairs under normal, uncongested conditions are relatively high, ranging from .76 for the
Northwest-Southwest market pair to .94 for the Northern California-Southern California market pair. All market pair correlations under "normal", uncongested times are statistically distinguishable from zero at the 1% level. Price correlations are higher for regions nearer in distance and lower for regions further apart, consistent with the hypothesis that regions which are nearer in distance are bound by arbitrage constraints more often (because transmission costs are lower) than regions which are located further apart. In addition, parameter estimates are consistent with the hypothesis that regional supply and demand factors which increase the level of congestion on the network decrease the extent to which the five regions in the WSCC are integrated. First, price correlations are lower in the Spring months, consistent with the hypothesis that regions are less tightly linked during the spring flush. As expected, price correlations are lower during peak periods, consistent with the hypothesis that utilization of the transmission grid affects the extent to which regions are integrated. All variables which explicitly indicate congestion on the network have the correct sign, though only one, congestion along the transmission line from the Northwest into the Southwest, is statistically significant. A low capacity rating on the southbound Pacific Intertie lowers price correlations as expected and is statistically significant. Finally, high water flows in the Northwest decrease price correlations, though the coefficient is not precisely estimated.

These results suggest that under normal, uncongested periods, prices in the five sub-regions in the WSCC are closely related, but that regional supply and demand conditions which give rise to congestion on the network cause prices to be less closely linked to one another. This latter result should be interpreted as implying that the relevant geographic market is more likely to narrow when these supply and demand conditions prevail. As with the standard price correlation test for
the geographic expanse of the market, this methodology does not provide concrete guidance for which levels of price correlations are high enough to definitively consider a set of regions integrated. The next section provides a more precise method to assess the expanse of the geographic market by examining the effect of congestion on arbitrage constraints.

1.5.2 The Effect of Congestion on Arbitrage Constraints

When the price of transmission service is regulated by a price cap, as it is in the wholesale electricity market in the WSCC, arbitrage constraints can be used to identify the effect of congestion on the geographic expanse of the market. In this Section, an empirical method is developed which assesses whether arbitrage constraints are binding and how transmission line congestion affects these arbitrage constraints. The interpretation of this technique is straightforward. If arbitrage constraints bind the wholesale prices in two regions together, the two regions are said to be in the same market. If arbitrage constraints do not bind the wholesale prices in the two regions together, the regions are said to be in different markets. The larger number of regions whose prices are all bound together by binding arbitrage constraints, the wider the expanse of the geographic market.

This approach estimates a variation of an arbitrage equation: \( P_A = \beta_0 + \beta_1 P_B + \epsilon \), where the flow of trade is from region B to region A and \( P_A \) and \( P_B \) are the prices in region A and B respectively. While the approach is similar in spirit to Horowitz (1981), this method incorporates known physical constraints to trades, changes in the transportation cost due to seasonal variation, losses
during transportation, and attention to the direction of flow of trade thereby extending the Horowitz methodology.

A further extension of Horowitz and subsequent literature based on this technique is to recognize the contemporaneous correlation of disturbances across regions in a network. In the WSCC, disturbances in the arbitrage equations include factors that are common to all of the regions, such as FERC announcements about merger approvals or denials as well as changes in the skills and technologies available to the regional electricity control operators. In addition, there are likely other omitted variables, such as unplanned generation plant outages and unplanned transmission line outages, that are common to all regions which cause the disturbances across equations to be correlated. Given the correlation of disturbances among regions in a network, the arbitrage equations are estimated jointly by seemingly unrelated regression (SUR) techniques.28

Another natural extension would seem to be to instrument for the price in region B since \( P_B \) and \( \varepsilon \) should be correlated: A positive shock to the price in region A should increases the demand for electricity purchased from region B thus raising the price in region B. This relationship though is exactly what a binding arbitrage equation implies and therefore what the arbitrage equation is testing: that shocks to one region are fully incorporated into the prices in the other

---

28Horowitz (1981) and variations on this approach in the literature use OLS to examine price relationships between pairs of prices.
region. For this reason, there are no instruments available if the arbitrage equation binds. Every factor correlated with $P_B$ is also correlated with $\varepsilon$.\(^2\)

The form of the arbitrage constraint for each market pair in a particular flow pattern (i.e. electricity flows from region 1 into region 2 and from region 1 into region 3) is:

$$P_A = \beta_0 + \beta_1 \text{PEAK} + \beta_2 P_B + \sum_{j=1}^{J} \gamma_j \text{CONGESTION}_j + \varepsilon$$

where regions A and B compose the market pair (region A is the buyer, region B is the seller), the dependent variable, $P_A$, is the price on date $t$ in region A, PEAK =1 if the observation comes from a peak period, $P_B$ is the price on date $t$ in region B, and CONGESTION$_j$=1 if there is a report of congestion on transmission line $j$ on date $t$. Each flow pattern yields an $M$ equation SUR model, where $M$ is the number of market pairs in the flow pattern. Each $m=1,2...,M$ arbitrage equation to be estimated for a particular flow pattern has $N$ observations. Stacking the $M$ arbitrage equations gives a total of $M*N$ observations used to estimate the system.

Expectations of the parameter estimates are straightforward. Electricity, like other goods which are transported for trade, experiences some degree of loss during transportation. For high voltage transfer of electricity, these "line losses" are typically 5% of electricity purchased. Assuming an

---

\(^2\)If $P_B$ and $\varepsilon$ are positively correlated, which they will be if the price for transmission service is a constant and arbitrage constraints bind, then parameter estimates from OLS (and SUR) are biased toward zero. Later sections find some evidence toward this end. Estimates of the mean price for transmission service in the next Section are slightly larger than estimates of the mean price for transmission service in the present Section.
5% line loss, this means for every one MW of electricity that region A purchases from region B only .95 MW arrive in region A. Therefore, for region A to receive 1 MW of electricity from region B it must purchase 1.053 (100/95) MW of electricity from region B. Consequently, if arbitrage constraints are binding between region A and region B, \( \hat{\beta}_2 \) is expected to be greater than 1 and about 1.053 assuming a line loss of 5%. Because line losses increase as the transmission distance increases, regions located closer together are expected to have a slightly lower coefficient on \( \hat{\beta}_2 \) while regions located further apart are expected to have a coefficient slightly greater than 1.053. If autarky prevails between the two regions, no relation is expected between the two prices and therefore \( \hat{\beta}_2 = 0 \) is expected.

If arbitrage constraints are binding, \( \hat{\alpha}_b = T > 0 \) is expected, where T is the transmission price in an off-peak period. The price of transmission service should increase in high electricity demand periods (the peak period) compared to low electricity demand periods (the off-peak period) because peak periods are generally a time of high wholesale electricity demand and therefore high demand for transmission. As a result, PEAK should enter positively. The price of transmission should increase as the distance between region A and region B increases. Looking across market pairs, it is expected that transportation costs in both peak and off-peak periods are smaller for regions located closer together than for regions located further apart. Since the transmission price for wholesale electricity is regulated by FERC, it is expected that both \( \hat{\alpha}_b \) and \( \hat{\beta}_o + \hat{\beta}_t \) are no greater than the regulated price cap on transmission service.
Congestion is expected to enter positively when the congestion is on the line directly connecting the two regions in the market pair because congestion increases the marginal cost of transmission along that line. Because the cap on the price of transmission set by FERC prevents the price of transmission from rising as high as the price for transmission service which would cause the arbitrage constraint to bind, it is expected that \( \hat{\beta}_b + \hat{\beta}_i + \gamma_j \) is greater than or equal to the regulated price cap.\(^{30}\) Line congestion has different effects depending on the line constrained and the market pair in question. In general, the constraint on the line is expected to have a positive effect on the price difference for market pairs located near a congested transmission line. For regions far from the congested transmission line, the constraint is expected to have a smaller positive effect, or no effect at all.

This technique is applied to the set of prices such that electricity flows from the Northwest into Southern California, from Northern California into Southern California, and from Southern California into the Inland Southwest. This pattern is observed most frequently in the data, thirty percent of the 948 sets of price observations follow this pattern, and is most commonly observed in the peak periods of Spring and Summer months.

There are four arbitrage equations to be estimated: Northwest→Northern California, Northwest→Southern California, Northern California→Southern California, and Southern California→Inland SW. The supplemental data indicate that congestion occurs on the lines

\(^{30}\) If the price of transmission service is not regulated, the competitive price for transmission service will rise to clear the market and arbitrage constraints will continue to bind. Even so, prices in the two regions will separate when congestion is present: the price in the low price region will not react to a positive shock in the high price region.
from the Northwest into Northern California (NWNCAF), on the line from the Northwest into Southern California (NWSCAF), on the line from Northern California into Southern California (NCASCAF), and on the line from Southern California into the Southwest (SCASWF). For each market pair, the arbitrage equation estimated is:

\[ P_A = \beta_0 + \beta_1 \text{PEAK} + \beta_2 P_B + \gamma_1 \text{NWNCAF} + \gamma_2 \text{NWSCAF} + \gamma_3 \text{NCASCAF} + \gamma_4 \text{SCASWF} + \epsilon \]

Each of the four (there are four market pairs) arbitrage equations contains 322 observations; a total of 1288 observations are used to estimate the SUR equation.\(^{31}\) Table 4 presents the generalized regression parameter estimates.

Coefficient estimates presented in Table 3 strongly suggest that the prices in the sub-regions of the WSRC are bound together by arbitrage constraints, but that congestion gives rise to price separation. Coefficient estimates for \(\hat{\beta}_2\) are as expected for all market pairs: Slightly greater than 1 assuming a reasonable amount of line loss. Table 4 illustrates that line loss increases as the transmission distance increases (2.06% for Northern California to Southern California - the nearest pair - and up to 7.5% for transmission from the Northwest into Southern California - the furthest pair). Estimated prices for off-peak transmission service are positive, significant, and less than the regulated price cap. Comparing across pairs of regions, off-peak prices for transmission service increase with distance from 0.431 Mills/KW from Northern California to

\(^{31}\) 322 of the 948 sets of price observations follow this directional electricity flow. Since there are 4 arbitrage equations to be estimated, each with 322 observations, when the 4 arbitrage equations are stacked to form the SUR system equation, there are a total of 322*4=1288 observations used in the estimation.
Southern California up to 1.045 Mills/KW from the Northwest to Southern California. In addition, the estimated price of transmission service from the Northwest to Northern California plus the estimated price of transmission service from the Northern California to Southern California equals the estimated price of transmission service from the Northwest to Southern California. For three pairs, the estimated increase in the price for transmission service in a peak period compared to an off-peak period is positive as expected although only precisely estimated for the Southern California-Southwest pair. For one pair, Northern California-Southern California, the estimated difference in transmission price in a peak period compared to an off-peak period is negative, but it is imprecisely estimated. One explanation for why the price of transmission service does not appear to vary between peak and off-peak periods is that in the WSCC the marginal cost of transmission service is relatively constant. The marginal cost of transmission service is relatively constant in the WSCC because reliability, rather than thermal constraints, limit transfers along the major transmission paths in the WSCC. For this reason, the marginal cost of transmission service, and therefore the price of transmission service, is relatively flat up to the rated limit.

Coefficients on congestion variables are generally positive and significant. For the Northwest-Northern California market pair, congestion on the transmission line directly connecting the regions (NWNCAF), causes a significant positive increase in the price difference by .83 Mills/Kwh. Parameter estimates for congestion on the two near transmission lines are not significant. The positive and significant coefficient on the SCASWF variable captures the fact that when there is high demand for imported wholesale electricity in the Southwest, the increased demand for wholesale electricity increases the demand for wholesale electricity generated in the
Northwest. Because the capacity of the transmission line from the Northwest into Northern California, one transmission route from the Northwest into the Southwest, is limited, the market pair Northwest-Northern California exhibits price separation.

For the Northwest-Southern California market pair, parameter estimates for congestion along direct and far transmission lines are not significant. Congestion on the near line (SCASWF) causes an increase in the price difference between the price in the Northwest and Southern California. The positive and significant coefficient on the SCASWF variable again captures the fact that when there is high demand for imported wholesale electricity in the Southwest, the increased demand for wholesale electricity increases the demand for wholesale electricity generated in the Northwest. The limited capacity of the transmission line from the Northwest into Southern California, another transmission route from the Northwest into the Southwest, causes the market pair Northwest-Southern California to exhibit price separation.

Parameter estimates in the Northern California-Southern California market pair indicate that congestion on the near transmission line (NWSCAF) causes an increase in the price difference between the price in Southern California and the price in Northern California. Congestion on the near transmission line (NWNCAF) and on the direct transmission line (NCASCA) cause a decrease in the price difference between Northern California and Southern California. This narrowing of the price difference when there is congestion along the transmission line connecting the Northwest to Northern California and the transmission line connecting Northern California to Southern California may be explained by energy flowing back up through Midway, the transmission juncture which divides California into a "Northern" and "Southern" half, from
Sylmar, the terminus of the DC line connecting the Northwest with Southern California. This feature causes the price in Southern California to fall relative to the price in Northern California thereby narrowing the price difference between Southern California and Northern California. Variables which indicate a report of congestion on the transmission grid are not precisely estimated in the Southern California - Inland Southwest market pair.

The results in this Section indicate that arbitrage constraints are binding across theWSCC under normal, uncongested conditions, but that congestion causes arbitrage constraints not to bind. In the next Section an empirical technique is developed to assess how frequently arbitrage constraints bind throughout theWSCC and how frequently prices in some regions become separated from prices in other regions.

1.5.3 How Frequently Do Arbitrage Constraints Bind?

Although results from the previous two Sections suggest that arbitrage constraints do bind and that congestion does cause price separation, these two Sections did not provide insight into how frequently congestion occurs nor how frequently wholesale prices in theWSCC are linked by arbitrage constraints. How frequently wholesale prices in theWSCC are linked by arbitrage constraints provides insight into how often the geographic expanse of the market is significantly smaller than the entire expanse of theWSCC. The method developed in this Section finds that arbitrage constraints bind prices across theWSCC 80% of the time, 19% of the time congestion arises which leads to price separations, and 1% of the time autarky prevails and the price difference between the regions is less than the price for transmission service.
The technique developed is similar to that proposed by Spiller-Huang (1986) and Spiller-Wood (1988), but expands on their method because this technique allows for three rather than two possible regimes: autarky, arbitrage, and congestion. Spiller-Huang and Spiller-Wood consider only autarky and arbitrage. In addition, the Spiller-Huang technique assumes the price of transmission service is exogenous. The price of transmission service, and transportation services in general, is not likely to be exogenous. For example, the mean price for transmission service, $T$, is expected to be higher in peak periods compared to off-peak periods ($T_{\text{peak}} > T_{\text{off-peak}}$) since the demand for transmission service is higher in peak periods than in off-peak periods. In addition, the probability of the three regimes are expected to be different under different prevailing supply and demand conditions. The congestion regime is expected to prevail more frequently in peak periods than in off-peak periods because high demand for electricity (peak period) increases utilization of the grid, and thus increases the likelihood of congestion occurring ($\lambda_{\text{Congestion, peak}} > \lambda_{\text{Congestion, off-peak}}$). The autarky regime is expected to prevail more frequently in off-peak periods than in peak periods because minimum load conditions arise in off-peak periods ($\lambda_{\text{Autarky, peak}} < \lambda_{\text{Autarky, off-peak}}$). Finally, the arbitrage regime is expected to prevail more frequently in off-peak periods than in peak periods because the transmission grid is not fully loaded in off-peak periods ($\lambda_{\text{Arbitrage, peak}} < \lambda_{\text{Arbitrage, off-peak}}$). The technique developed in this Section allows for transaction costs and the probabilities of the three states - arbitrage, autarky, and congestion - to be different under different prevailing supply and demand conditions.

i. Technique
Consider two regions: A and B. In period $t$, if both regions fill retail electricity demand using only own generating capacity, the price for wholesale electricity in each region is the autarky price: $P^A_{t, autarky}$ and $P^B_{t, autarky}$. Assume $P^A_{t, autarky} > P^B_{t, autarky}$. Denote the price for transmission service from region B to region A in period $t$ as $T$, where $T = T + \varepsilon_t$ and $\varepsilon_t \sim N(0, \sigma^2_e)$. There are three mutually exclusive and exhaustive possibilities for the relationship between the prevailing prices in region A, $P^A_t$, and region B, $P^B_t$:

1) Autarky

If the autarkic prices differ by less than $T$, then the price for wholesale electricity which prevails in region A, $P^A_t$ and in region B, $P^B_t$, are the autarkic prices: $P^A_t = P^A_{t, autarky}$ and $P^B_t = P^B_{t, autarky}$:

$$ P^A_t - P^B_t - T < \varepsilon_t. \quad (eq. 1) $$

2) Arbitrage

If the autarkic prices differ by more than $T$, and there is no transmission line congestion, then the prevailing prices for wholesale electricity satisfy:

$$ P^A_t - P^B_t - T = \varepsilon_t. \quad (eq. 2) $$

3) Congestion

39
If the autarkic prices differ by more than $T$, and there is transmission line congestion, then the prevailing prices for wholesale electricity satisfy:

$$P_i^A - P_i^B - T > \varepsilon_i.$$  \hspace{1cm} (eq. 3)

The probability of observing each regime - autarky, arbitrage and congestion - is a constant: $\lambda^{\text{Autarky}}$, $\lambda^{\text{Arbitrage}} = 1 - \lambda^{\text{Autarky}} - \lambda^{\text{Congestion}}$, and $\lambda^{\text{Congestion}}$ respectively.

Define $\nu_i$ to be a positive random variable. Then the three mutually exclusive and exhaustive possibilities for the relationship between the prevailing prices in region A and region B are in fact a switching regressions system where,

$$P_i^A - P_i^B - T = \varepsilon_i - \nu_i$$ \hspace{1cm} (eq. 4)

occurs with probability $\lambda^{\text{Autarky}}$ (corresponding to the autarky regime);

$$P_i^A - P_i^B - T = \varepsilon_i$$ \hspace{1cm} (eq. 5)

occurs with probability $1 - \lambda^{\text{Autarky}} - \lambda^{\text{Congestion}}$ (corresponding to the arbitrage regime); and

$$P_i^A - P_i^B - T = \varepsilon_i + \nu_i$$ \hspace{1cm} (eq. 6)
occurs with probability $\lambda^\text{Congestion}$ (corresponding to the congestion regime).

In this paper, the positive error component is assumed to be distributed independently of $\varepsilon_i$, with a truncated (from below at zero) Normal distribution, $N(0, \sigma^2 \varepsilon)$. Denote

$$\theta = \{T, \sigma_\varepsilon, \sigma_\nu, \lambda^\text{Autarky}, \lambda^\text{Congestion}\}$$

as the parameter vector. The likelihood function for the N observations is given by:

$$L = \prod_{i=1}^{N} \left\{ \lambda^\text{Autarky} \cdot f_i^\text{Autarky} + (1 - \lambda^\text{Autarky} - \lambda^\text{Congestion}) \cdot f_i^\text{Arbitrage} + \lambda^\text{Congestion} \cdot f_i^\text{Congestion} \right\}$$

where $f_i^\text{Autarky}$, $f_i^\text{Arbitrage}$, and $f_i^\text{Congestion}$ are the density functions of (eq. 4), (eq. 5), and (eq. 6) respectively. Appendix 1 derives the form of the density functions. The maximum likelihood estimates are obtained by maximizing the logarithm of the likelihood function with respect to $\theta$.

ii. Estimation and Interpretation

As with the previous methodologies, it is important to apply this technique taking account of the flow of electricity and prevailing supply and demand conditions. This technique is applied to the Northwest and Southern California regions with electricity flowing in a southbound direction (from the Northwest into Southern California). Price data matching this flow criterion are
partitioned into two categories: peak period observations and off-peak period observations.\textsuperscript{32}

Figure 8 presents a histogram of the price differences for the Southern California - Northwest pair when electricity moves in the southbound direction. That the distribution of the price differences in peak periods compared to off-peak periods are different suggests that transaction costs and the probabilities of the three regimes are not exogenous, and instead depend on prevailing supply and demand conditions.

Table 4a presents the parameter estimates obtained from maximum likelihood estimation using peak period observations. All of the parameter estimates are of the expected sign and reasonable magnitude. All but one parameter estimate, the probability of observing the autarky regime, are statistically distinguishable from zero at the 5% level. The estimated mean price for transmission service between Southern California and the Northwest, $T$, is 1.99 Mills/Kwh, which agrees with the estimated price for transmission service in Table 3. The probabilities of the three regimes in peak periods are reasonable and agree with both anecdotal evidence\textsuperscript{33} from the WSCC as well as with the flavor of the earlier sections in this paper: 1% of the time autarky prevails, efficient arbitrage prevails 80% of the time, and 19% of the time congestion prevents efficient arbitrage.

Table 4b presents the parameter estimates obtained from maximum likelihood estimation using off-peak period observations. All of the coefficients are of the correct sign and reasonable magnitude, but only three parameter estimates are statistically distinguishable from zero at the

\textsuperscript{32}There are not sufficient data to partition the observations more finely than peak and off-peak. Maximum likelihood techniques depends on a relatively large number of observations for estimation to proceed reliably.

\textsuperscript{33}See Lehr and VanVactor (1997) and McCullough (1996).
5% level. The estimated mean price for transmission service between Southern California and the Northwest, \( T \), is 1.66 Mills/Kwh and agrees with the estimated price for transmission service in Table 3. The probabilities of the three regimes occurring in off-peak periods are reasonable: 9% of the time autarky prevails, efficient arbitrage prevails 83% of the time, and 8% of the time congestion prevents efficient arbitrage. Neither the probability of autarky nor the probability of congestion are statistically distinguishable from zero at the 5% level, implying that efficient arbitrage prevails 100% of the time in off-peak periods.

Estimates of the level of the probabilities for the three regimes compared across peak and off-peak periods are as expected. First, results indicate that efficient arbitrage is slightly more likely to occur in off-peak periods compared to peak periods, \( \lambda_{\text{Arbitrage}}^{\text{Peak}} = 80\% < 83\% = \lambda_{\text{Arbitrage}}^{\text{Off-Peak}} \).

Second, results indicate that price separation is more likely to occur due to congestion in peak periods compared to off-peak periods, \( \lambda_{\text{Congestion}}^{\text{Peak}} = 19\% > 8\% = \lambda_{\text{Congestion}}^{\text{Off-Peak}} \). Third, results indicate that autarky is more likely to occur in off-peak periods than in peak periods, \( \lambda_{\text{Autarky}}^{\text{Peak}} = 1\% < 9\% = \lambda_{\text{Autarky}}^{\text{Off-Peak}} \). Finally, comparing parameter estimates of the mean price for transmission service between peak and off-peak observations, the estimated mean price for transmission service is lower in off-peak periods compared to peak periods, which agrees with both expectations as well as with estimates from Table 3.

While the previous two Sections indicate that congestion occurs on transmission lines in the WSCC and that congestion causes regions to exhibit price separation, these Sections did not provide evidence of how frequently arbitrage occurs nor how frequently prices in the WSCC
regions become separated. The technique developed in this Section demonstrates that for the Northwest-Southern California market pair, arbitrage constraints bind approximately four-fifths of the observations. These results indicate that, by and large, the geographic expanse of the wholesale electricity market extends across the WSCC so that under normal uncongested conditions, wholesale electric customers are able to turn to generation suppliers dispersed over a wide geographic area, allowing them to purchase or sell electricity either within their control areas, in directly connected control areas, or in more remote control areas. On the other hand, in nineteen percent of the observations, price separation occurs. In these times, the geographic expanse of the market narrows and the effective number of competing firms falls.

1.6 CONCLUSION AND IMPLICATIONS

This paper explored spatial dispersion of prices and the geographic expanse of a market in a dynamic setting. Spatial price relationships can be used as a means to infer the geographic expanse of the market, and more generally, market performance. In particular, the larger the geographic market, the more competitors, and therefore the less likely firms will be able to exercise market power. This paper suggests that the geographic expanse of the market for wholesale electricity is quite wide under most conditions and in most time periods, but that imperfect competition may arise as a result of transmission line congestion. Although a "narrow" geographic market may be more conducive to market power than a "wide" geographic market, a smaller geographic market does not immediately imply that market power problems will necessarily arise in these times. A smaller geographic market may still have enough competitors in generation services to alleviate market power concerns.
The behavior of the current wholesale electricity market can also go a long way toward informing the discussion of pricing behavior and performance in a restructured electricity industry. One important input into a complete assessment of imperfect competition in a restructured electricity industry is the geographic expanse of the market for generation services. This paper suggests that the geographic expanse of the market for retail electricity in a restructured electricity market is likely to be quite wide, and therefore relatively competitive, under most conditions and in most time periods, but that transmission line congestion can cause the geographic expanse of the market to narrow giving rise to imperfect competition. Finally, the techniques developed in this paper can readily be applied to data once a restructured electricity market emerges. Because market conditions in the electricity industry are likely to change significantly in the next few years as the structure of the electricity sector changes dramatically, the present analysis will be a useful benchmark against which to compare post-restructuring wholesale price relationships.
REFERENCES


Western Regional Transmission Association. *WRTA Member Tariff Summaries.*


APPENDIX 1

The density functions $f_{i, \text{Autarky}}$, $f_{i, \text{Arbitrage}}$, and $f_{i, \text{Congestion}}$ are derived as follows. Denote $\phi$ and $\Phi$ as the standard Normal density and cumulative distribution functions respectively. The error term in the Autarky and Congestion equations is a composite error term, composed of one term distributed as a Normal ($\varepsilon_i \sim N(0, \sigma_{\varepsilon}^2)$) and one term distributed as a truncated Normal ($\nu_i \sim N(0, \sigma_{\nu}^2)$ with $\nu_i > 0$).

The density function for the autarky regime is given by:

$$
\varepsilon_i - \nu_i = P_i^A - P_i^B - T
$$

$$
f_{i, \text{Autarky}} = f(\varepsilon_i - \nu_i) = \frac{2}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\nu}^2}} \cdot \phi \left( \frac{P_i^A - P_i^B - T}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\nu}^2}} \right) \cdot \left[ 1 - \Phi \left( \frac{(P_i^A - P_i^B - T)(\frac{\sigma_{\nu}}{\sigma_{\varepsilon}})}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\nu}^2}} \right) \right]
$$

The density function for the arbitrage regime is given by:

$$
\varepsilon_i = P_i^A - P_i^B - T
$$

$$
f_{i, \text{Arbitrage}} = f(\varepsilon_i) = \frac{1}{\sigma_{\varepsilon}} \cdot \phi \left( \frac{P_i^A - P_i^B - T}{\sigma_{\varepsilon}} \right)
$$

The density function for the congestion regime is given by:

$$
\varepsilon_i + \nu_i = P_i^A - P_i^B - T
$$

$$
f_{i, \text{Congestion}} = f(\varepsilon_i + \nu_i) = \frac{2}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\nu}^2}} \cdot \phi \left( \frac{P_i^A - P_i^B - T}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\nu}^2}} \right) \cdot \left[ 1 - \Phi \left( \frac{-(P_i^A - P_i^B - T)(\frac{\sigma_{\nu}}{\sigma_{\varepsilon}})}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\nu}^2}} \right) \right]
$$
FIGURE 1
The WSCC Region
FIGURE 2
Major Transmission Paths in theWSCC
FIGURE 3a
Rating of the Southbound AC Pacific Intertie
June 23 1995 - May 9 1997

Maximum Southbound Capacity of the Pacific
AC Intertie: 4800 MW
FIGURE 3b
Rating of the Southbound DC Pacific Intertie
June 23 1995 - May 9 1997

Maximum Southbound Capacity of the Pacific DC Intertie: 3100 MW
FIGURE 4a
Peak Pre-Scheduled Wholesale Electricity Movements in the WSCC
June 23 1995 - December 31 1996
FIGURE 4b
Off-Peak Pre-Scheduled Wholesale Electricity Movements in the WSCC
June 23 1995 - December 31 1996
FIGURE 5
Wholesale Electricity Prices (Mills/Kwh) Under Typical Electricity Flows

Figure 4a
Spring, Daytime

10.9

NW

12

SCA

11.1

SW

Figure 4b
Spring, Nighttime

5.1

NW

6.1

SCA

6.8

SW

Figure 4c
Fall-Winter, Nighttime

17.0

NW

16.4

SCA

14.8

SW

55
FIGURE 6
Price Relationships: Heat Wave in Southwest
FIGURE 7a
Wholesale Electricity Trade
Competitive Generation and Competitive Transmission Service.
No Constraints to Trade

Marginal Cost

competitive price
for transmission service

Marginal Cost

W X

0 QA Quantity of Electricity Qb

Region A → Region B

MC_A

MC_B

0

Q_A retail electricity demand in region A
Q_B retail electricity demand in region B
W quantity of electricity generated by Utility A to fulfill retail demand in region A
Z quantity of electricity generated by Utility B to fulfill retail demand in region B
X quantity of electricity sold by Utility B to Utility A

Π_A gain from wholesale trade in region A
Π_B gain from wholesale trade in region B
FIGURE 7b
Wholesale Electricity Trade: Increase Demand in Region A
Competitive Generation and Competitive Transmission Service.
No Constraints to Trade

Q_A' retail electricity demand in region A
Q_B retail electricity demand in region B
W' quantity of electricity generated by Utility A to fulfill retail demand in region A
Z quantity of electricity generated by Utility B to fulfill retail demand in region B
X' quantity of electricity sold by Utility B to Utility A
FIGURE 7c
Wholesale Electricity Trade: Increase Demand in Region A
Competitive Generation and Competitive Transmission Service.
Constraint to Trade

Q'_A  retail electricity demand in region A
Q_B  retail electricity demand in region B
W''  quantity of electricity generated by Utility A to fulfill retail demand in region A
Z  quantity of electricity generated by Utility B to fulfill retail demand in region B
X  quantity of electricity sold by Utility B to Utility A
FIGURE 8
Price Difference = P_{SCA} - P_{NW}

Frequency of Price Difference: Off-Peak

Frequency of Price Difference: Peak
### TABLE 1
Sources of Non-Firm Imported Energy, 1995

**GWhs Purchased from Region***

<table>
<thead>
<tr>
<th></th>
<th>Pacific Northwest</th>
<th>Northern California</th>
<th>Southern California</th>
<th>Inland Southwest</th>
<th>Central Rockies</th>
<th>Canada/Mexico</th>
<th>Power Marketer</th>
<th>All Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacific Northwest</td>
<td>51,554</td>
<td>3,341</td>
<td>1,664</td>
<td>714</td>
<td>1,108</td>
<td>2,875</td>
<td>4,862</td>
<td>66,118</td>
</tr>
<tr>
<td>Northern California</td>
<td>1,323</td>
<td>8,676</td>
<td>880</td>
<td>114</td>
<td>396</td>
<td>296</td>
<td>310</td>
<td>11,995</td>
</tr>
<tr>
<td>Southern California</td>
<td>4,529</td>
<td>1,740</td>
<td>1,028</td>
<td>2,265</td>
<td>1,416</td>
<td>413</td>
<td>1,335</td>
<td>12,726</td>
</tr>
<tr>
<td>Inland Southwest</td>
<td>1,074</td>
<td>165</td>
<td>312</td>
<td>1,534</td>
<td>225</td>
<td>0</td>
<td>689</td>
<td>3,999</td>
</tr>
<tr>
<td>Central Rockies</td>
<td>1,567</td>
<td>1,292</td>
<td>219</td>
<td>1,050</td>
<td>3,216</td>
<td>85</td>
<td>4,361</td>
<td>11,790</td>
</tr>
</tbody>
</table>

*Data was obtained from Energy Information Agency, *Wholesale Electric (Bulk Power) Trade Data*. Purchases from Qualifying Facilities have been excluded.
### TABLE 2
The Effect of Regional Supply and Demand Conditions: OLS Parameter Estimates

\[ \rho_i = \sum_{k=1}^{10} \beta_k \text{PAIR}_{i,k} + \beta_1 \text{PEAK}_i + \beta_2 \text{SPRING}_i + \beta_3 \text{SUMMER}_i + \beta_4 \text{WINTER}_i + \beta_5 \text{NWNCADF}_i + \beta_6 \text{WSCAF}_i + \beta_7 \text{CJDAMH}_i + \beta_8 \text{DERATING}_i + \epsilon_i \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient</th>
<th>White Corrected Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Pair:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NW, NCA</td>
<td>0.925***</td>
<td>0.045</td>
</tr>
<tr>
<td>NW, SCA</td>
<td>0.900***</td>
<td>0.043</td>
</tr>
<tr>
<td>NW, SW</td>
<td>0.763***</td>
<td>0.055</td>
</tr>
<tr>
<td>NW, CR</td>
<td>0.835***</td>
<td>0.050</td>
</tr>
<tr>
<td>NCA, SCA</td>
<td>0.937***</td>
<td>0.043</td>
</tr>
<tr>
<td>NCA, SW</td>
<td>0.791***</td>
<td>0.058</td>
</tr>
<tr>
<td>NCA, CR</td>
<td>0.856***</td>
<td>0.050</td>
</tr>
<tr>
<td>SCA, SW</td>
<td>0.795***</td>
<td>0.063</td>
</tr>
<tr>
<td>SCA, CR</td>
<td>0.853***</td>
<td>0.051</td>
</tr>
<tr>
<td>SW, CR</td>
<td>0.846***</td>
<td>0.056</td>
</tr>
<tr>
<td>SPRING</td>
<td>-0.074</td>
<td>0.061</td>
</tr>
<tr>
<td>SUMMER</td>
<td>0.106***</td>
<td>0.038</td>
</tr>
<tr>
<td>WINTER</td>
<td>0.087**</td>
<td>0.040</td>
</tr>
<tr>
<td>PEAK</td>
<td>-0.066***</td>
<td>0.021</td>
</tr>
<tr>
<td>NWNCADF</td>
<td>-0.064</td>
<td>0.043</td>
</tr>
<tr>
<td>NWSCAF</td>
<td>-0.051*</td>
<td>0.030</td>
</tr>
<tr>
<td>DERATING</td>
<td>-0.042*</td>
<td>0.022</td>
</tr>
<tr>
<td>CJDAMH</td>
<td>-0.014</td>
<td>0.022</td>
</tr>
</tbody>
</table>

N=400; R^2=.93

***significant at 1%; **significant at 5%; *significant at the 10% level.
### TABLE 3
Estimating the Effect of Congestion on Arbitrage Constraints: SUR Parameter Estimates
(standard errors in parentheses)

\[ P_s = \beta_0 + \beta_1 \text{PEAK} + \beta_2 P_a + \gamma_1 NWNCAF + \gamma_2 NWSCAF + \gamma_3 NCASCAF + \gamma_4 SCASWF + \epsilon \]

<table>
<thead>
<tr>
<th></th>
<th>NW-&gt;NCA</th>
<th>NW-&gt;SCA</th>
<th>NCA-&gt;SCA</th>
<th>SCA-&gt;SW</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{NW} )</td>
<td>1.058***</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( P_{NW} )</td>
<td>—</td>
<td>1.081***</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.021)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( P_{NCA} )</td>
<td>—</td>
<td>—</td>
<td>1.021***</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>( P_{SCA} )</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>1.044***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.015)</td>
</tr>
<tr>
<td>( T_{NW-&gt;NCA} )</td>
<td>0.602***</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.180)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T_{NW-&gt;SCA} )</td>
<td>—</td>
<td>1.045***</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.195)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T_{NCA-&gt;SCA} )</td>
<td>—</td>
<td>—</td>
<td>0.431***</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.097)</td>
<td></td>
</tr>
<tr>
<td>( T_{SCA-&gt;SW} )</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.731***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.159)</td>
</tr>
<tr>
<td>PEAK</td>
<td>0.184</td>
<td>0.025</td>
<td>-0.161</td>
<td>0.335**</td>
</tr>
<tr>
<td></td>
<td>(0.186)</td>
<td>(0.210)</td>
<td>(0.132)</td>
<td>(0.155)</td>
</tr>
<tr>
<td>NWNCAF</td>
<td>0.833***</td>
<td>0.339</td>
<td>-0.512**</td>
<td>0.172</td>
</tr>
<tr>
<td></td>
<td>(0.235)</td>
<td>(0.280)</td>
<td>(0.208)</td>
<td>(0.202)</td>
</tr>
<tr>
<td>NWSCAF</td>
<td>-0.306</td>
<td>0.318</td>
<td>0.630***</td>
<td>-0.227</td>
</tr>
<tr>
<td></td>
<td>(0.280)</td>
<td>(0.333)</td>
<td>(0.246)</td>
<td>(0.239)</td>
</tr>
<tr>
<td>NCASCAF</td>
<td>1.019*</td>
<td>-0.495</td>
<td>-1.536***</td>
<td>-0.108</td>
</tr>
<tr>
<td></td>
<td>(0.604)</td>
<td>(0.719)</td>
<td>(0.534)</td>
<td>(0.517)</td>
</tr>
<tr>
<td>SCASWF</td>
<td>2.650***</td>
<td>3.116***</td>
<td>0.409</td>
<td>0.197</td>
</tr>
<tr>
<td></td>
<td>(0.647)</td>
<td>(0.770)</td>
<td>(0.571)</td>
<td>(0.561)</td>
</tr>
</tbody>
</table>

\( N=322 \quad R^2=.93 \)  \( N=322 \quad R^2=.91 \)  \( N=322 \quad R^2=.95 \)  \( N=322 \quad R^2=.96 \)

***significant at 1%, **significant at 5%, *significant at 10%
**TABLE 4a**
Frequency Arbitrage Constraints Bind: MLE Parameter Estimates
Electricity flows from the Northwest into Southern California, Peak

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>1.993***</td>
<td>0.569</td>
</tr>
<tr>
<td>$\sigma_e$</td>
<td>1.531***</td>
<td>0.518</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>1.711***</td>
<td>0.296</td>
</tr>
<tr>
<td>$\lambda^A$ Autarky</td>
<td>0.011</td>
<td>0.596</td>
</tr>
<tr>
<td>$\lambda^C$ Congestion</td>
<td>0.198**</td>
<td>0.106</td>
</tr>
</tbody>
</table>

N=146.
***significant at 1%, **significant at 5%, *significant at 10%

---

**TABLE 4b**
Frequency Arbitrage Constraints Bind: MLE Parameter Estimates
Electricity flows from the Northwest into Southern California, Off-Peak

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>1.668***</td>
<td>0.548</td>
</tr>
<tr>
<td>$\sigma_e$</td>
<td>0.741***</td>
<td>0.239</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>0.490***</td>
<td>0.167</td>
</tr>
<tr>
<td>$\lambda^A$ Autarky</td>
<td>0.094</td>
<td>1.635</td>
</tr>
<tr>
<td>$\lambda^C$ Congestion</td>
<td>0.084</td>
<td>0.352</td>
</tr>
</tbody>
</table>

N=161.
***significant at 1%, **significant at 5%, *significant at 10%
CHAPTER 2

Allowance Trading Activity and State Regulatory Rulings: Evidence from the U.S. Acid Rain Program

2.1 INTRODUCTION

The U.S. Acid Rain Program was created by Title IV of the 1990 Clean Air Act Amendments (1990 CAAA). It is one of the first, and by far the most extensive, applications of a market based approach to pollution control. The Acid Rain Program applies to fossil fuel fired electric utilities in the 48 contiguous United States requiring the affected generating units to achieve, in sum, a reduction of 10 million tons of sulfur dioxide (SO$_2$) per year from 1980 emissions levels and a 2 million ton reduction of nitrogen oxide (NO$_X$) emissions annually by the year 2000. Sulfur dioxide and nitrogen oxide are thought to be the two primary precursors of acid rain. The Acid Rain Program was implemented in two Phases, with the first round of emission limitations taking effect in 1995, and a stricter round of emissions limitations taking effect in the year 2000.

Compared to the historical "command and control" approach to pollution abatement, the market based approach to acid rain control instituted by the 1990 CAAA recognizes that the cost of pollution abatement is not identical across all generating units. Tradable permits, called allowances, are allocated to affected electric utility generating units. Each allowance represents one ton of SO$_2$. The firms are free to buy and sell the allowances with few restrictions in order to reduce aggregate SO$_2$ emissions at the least cost. In theory, the higher marginal cost of abatement units reduce emissions by less, purchasing allowances to cover their higher SO$_2$ emissions, while the lesser cost of abatement units reduce SO$_2$ emissions by more, selling
allowances generated from over-compliance to those units with higher abatement costs. Under Title IV, the only obligation on the part of the generating units is that at the end of every "true-up period" each affected generating unit must hold an allowance with a vintage year of that year or earlier for each ton of SO2 emitted in that year.

Earlier tradable permits programs, such as the Environmental Protection Agency's (EPA) emission trading program, which was initiated in 1974 to curb smog in the Los Angeles basin, have tended to have more restrictive trading guidelines than the EPA's Acid Rain Program. While the smog control program restricts the ability to bank permits, devalues banked permits, restricts trading between geographic regions, and requires regulatory approval for permit trades, the Acid Rain Program allows for inter-temporal trading ("banking"), unrestricted geographic trading, and does not require the EPA's approval of allowance trades. The only federal restriction on trading is that allowances may not be borrowed from future vintages for use in the current compliance year.

From the beginning, there has been concern whether utilities would choose to participate in allowance trading, and whether regulatory activity at the state level would further complicate utilities' decision to trade allowances. Several sources suggested early on that electric utilities would be reluctant to engage in inter-utility allowance trading activity for a variety of reasons.

---

1The "true-up" period is the thirty days (1 January to 30 January) following the compliance year during which affected units may do last minute buying or transferring of allowances into their unit accounts in order to match allowance holdings with their emissions tonnage for the year for which compliance is being established without incurring penalties. After these 30 days the affected units must surrender the appropriate number of allowances from their unit accounts to the EPA, on a first in first out basis unless otherwise requested, or be subject to a fine of $2,000 per ton. Gross non-compliance may result in criminal proceedings.

2See Foster and Hahn 1995.

3The differences in the freedom to trade permits in the two programs are due, in part, to differences in the nature of the pollutants, volatile organic chemicals, which cause smog, and the nature of the pollutant, sulfur dioxide, which causes acid rain.
including regulatory, industry, and market factors. This concern was magnified by a second, and inter-related concern, that misrepresentations by the popular press of the tradable permits program would cause state public utility commissions (PUC) to bow to pressure from local environmentalists and constituents, and balk at allowing utilities to trade allowances. Fullerton, McDermott, and Caulkins (1996) argue that the potential cost-increasing impact of state regulatory behavior could be substantial.

The ease with which utilities can engage in both internal and arms lengths exchanges will be an important determinant of the performance of the market, in particular, it may suggest reasons why the allowance trading program may (or may not) achieve all potential cost savings available in a market based approach as compared to a command and control approach to pollution control. In addition, it is important, as well as interesting, to understand which states have been trading allowances because it provides important evidence on the trading process itself. Finally, it is important to understand trading behavior in a tradable permits market in order to improve future market based approaches to pollution control.

The purpose, then, of this paper is to determine whether, and the extent to which, PUC rulings on allowance trading activity explain observed allowance trading. Because regulatory rulings may be endogenous, a simultaneous equations model is used to analyze the effect of state regulatory rulings on allowance trading activity. A reduced form model is used to assess the direct and indirect effects of regulatory and non-regulatory activity on allowance trading behavior. The results indicate that regulation has been conducive to allowance trading activity in the early years

---


5 See, for example, Wald 1995, a New York Times article which describes the Acid Rain Program as a Federal law which creates a national market to pollute.

6 See Passell 1996.
of the Acid Rain Program. The final Section adds several caveats to drawing hasty policy implications from this conclusion.

The organization of this paper is as follows. Section 2 sets the stage by providing a brief overview of allowance trading activity observed to date, discusses PUC regulations of emission allowances, and considers the endogeneity of regulations and the regulated activity. Section 3 describes the empirical framework and the source and scope of the data used, as well as its limitations. Section 4 presents empirical results. Section 5 assesses whether PUC regulations affect the extent to which allowance trading activity is undertaken. Section 6 offers concluding remarks.

2.2 ALLOWANCE TRADING ACTIVITY AND PUC REGULATIONS

A. Allowance Trading Activity

According to the EPA's Allowance Transaction System (ATS), the official record of allowance holdings for compliance purposes and the primary source for allowance trading data, it appears that over 32 million allowances have been traded since the program's inception, since that is the total number of allowances that have been recorded changing accounts.7 There are several reasons why this may be an inaccurate way of examining allowance trading activity. First, trades for allowances to be used for compliance in the current compliance year must be recorded in the ATS files, but it need not be earlier than the "true up" deadline. As a result, the ATS may lead to an underestimate of the number of trades occurring and the volume of allowances being traded. Consequently, the ATS data has been supplemented with data from reports in Energy Daily as well as from reports in Clean Air Compliance Review (CACR), a publication specifically targeted to issues of clean air compliance for stationary sources. A comparison of trades listed in the trade press to trades listed in the ATS files indicate that most trades have been registered with the ATS regardless of the allowances' vintage or the year the unit or operating company intends to

7 Approximately 268 million allowances have been issued and are available to trade.
use the allowances for compliance purposes. The ATS may also underestimate the number of trades occurring and the volume of allowances being traded because the ATS does not record options to buy or sell allowances that have not been exercised.

Second, it is important to distinguish between what trading activity is a trade in its proper sense and what activity is merely transfer activity. A trade is defined to be a considered decision to move emission allowances based in whole or in part on the price of allowances, the compliance strategies of the unit(s), if any, involved, or for strategic reasons. Trades include allowances bought in the EPA's annual auctions, allowances sold in the EPA's annual private auctions, and movements of allowances between plants, units, operating companies, brokers, fuel companies, individuals, and organizations for considered reasons rather than for reallocation, accounting, and/or joint ownership agreements. All other movements of allowances between two accounts, including reallocation, accounting, and/or prior contractual arrangements, are defined to be transfers. The distinction between a trade and a transfer is that a trade is an exchange that is considered in nature, such as for cost savings reasons, and is not simply the result of an accounting arrangement or a prior ownership agreement.

This paper will be concerned solely with arms-length allowance trades since it is difficult to disentangle the reason why an allowance may be transferred internally (e.g. economic versus accounting) without additional and usually propriety information. Internal -- intra-utility -- transactions include transactions between units within the same utility as well as between utilities within the same holding company. Arms-length -- inter-utility -- transactions are more clearly classified as motivated by economic reasons than intra-utility transactions. Including only executed, arms-length, trades involving at least one utility, approximately 4 million allowances have been traded since the program's inception.
B. Public Utility Commission Regulation

A utility's decision to trade allowances may be affected by the nature and behavior of its public utility commission. A utility's trading activity may be influenced by PUC behavior primarily because expenditure decisions made by a utility are subject to a prudence test by the utility's PUC. A prudence test determines whether a purchase or sale by a utility was reasonable under the circumstances that were known, or reasonably knowable, at the time of the expenditure or sale. For example, if a purchase is deemed reasonable, then cost recovery (usually through the rate base) of the expenditure is allowed; if a purchase is deemed imprudent, then the utility is not permitted to pass the costs incurred on to the ratepayers. Utilities may also perceive a risk associated with allowance sales. In particular, utilities may perceive a risk that during a ratemaking case, the price that any allowances were sold at will be questioned by the commission as to whether it was the "best" available price, or that they will be chastised by the commission for not seeking out a "better" price. Because emission allowances are a relatively new cost for a PUC to assess the prudence of, utilities may perceive an added risk when trading allowances, in particular, that the commission will be inexperienced in judging the prudence of allowance purchases, and will, therefore, too frequently determine allowance trades imprudent.

Formal PUC regulations, called generic orders, as well as informal PUC rulings, called guidelines, may mitigate a utility's perceived risk of trading emission allowances. Although a guideline does not carry the same force as an order, guidelines nevertheless convey the commission's attitude and intent.8 Generic orders as well as guidelines on allowance trading indicate the state PUC's expected treatment of emission allowances in a ratemaking case thereby minimizing the possibility that a utility's allowance trading activity will be ruled imprudent in a prudence review. It is possible that utilities in states with no formal ruling engaged in informal

---

8 The issuance of a guideline versus a generic order may in itself be endogenous: Frequent trading activity may be associated with the issuance of a formal order, while infrequent trading activity may be associated with the issuance of an informal order. This point is not dealt with in this paper.
conversations with their state commission regarding the ratemaking treatment of allowances, but this type of guidance is much less certain and certainly less secure.

As of the close of 1995, fifteen state public utility commissions had explicitly addressed the issue of allowance trading through the issuance of a formal generic order or an informal guideline. A complete discussion of the state regulatory rulings from the fifteen state PUCs which have addressed the issue of allowance regulation is reserved for the Appendix; Table 1 and Table 2 summarize information about the states whose PUC has issued a guideline or generic order. Table 1 summarizes the states whose PUC has issued a guideline or generic order on the regulatory treatment of allowances. There are several general observations about the regulatory treatment of allowances that have been issued, formally or informally, by state public utility commissions. First, the regulations largely require one hundred percent of both expenses and revenues to be returned to the ratepayers. In terms of accounting practices, the net gain (or loss) incurred from allowance transactions are used to offset (or increase) fuel costs. Second, a few states have taken an incentive based approach to allowances, allowing the utility to retain a portion of any gains from allowance sales beyond those sales which are below the line. 9 Finally, the regulations are often drawn to a state's specific circumstance. States with a large bank of allowances, due to a pre-existing state SO2 cap which makes units in those states largely unconstrained by Title IV (e.g. Wisconsin, Connecticut) or due to Phase 1 Extension Bonuses (e.g. Pennsylvania), have issued regulations creating favorable conditions for utilities to sell allowances, while states anticipating the purchase of allowances have issued regulations encouraging favorable conditions for utilities to buy allowances (e.g. Ohio, North Carolina). Table 2 illustrates the increase in number of states with guidelines or generic orders in effect. The number of states issuing regulatory statements on allowance trading has grown sharply over time: From zero in 1992, to ten in 1993, to fifteen by the close of 1995. As the number of states

---

9A below the line transaction is one that only involves shareholder monies. All gains or losses from below the line transactions are absorbed by the electric utility shareholders not the ratepayers. Thus far, all rulings permit utilities to retain one hundred percent of the gains and losses from below the line transactions.
issuing regulations on the treatment of allowances has increased, so too has the number of states with utilities engaging in allowance trading activity. From zero in 1991, the number of states with utilities engaging in inter-utility allowance trading grew to twelve in 1993, eighteen in 1994 and then to thirty in 1995, the first year that Phase 1 emission limitation requirements were in effect.\textsuperscript{10} Looking over all the years, thirty-one states have at least one utility that has engaged in allowance trading activity.

\textit{C. Endogeneity}

The effect of the issuance of regulations on allowance trading activity may be endogenous to the utility's decision to trade allowances. A state's PUC regulation on allowance trading activity may increase observed trading activity in that state by minimizing risk born by utilities in that state. In addition, increased allowance trading activity by utilities in a state may increase the likelihood that the state PUC will issue a regulation on allowance trading activity in order to address the issue of allowance trading comprehensively rather than on a case by case basis.

A thorough reading of the guidelines and generic orders issued by state PUCs, other PUC documents pertaining to the ratemaking treatment of allowances, as well as conversations with PUC staff directly involved in the regulatory treatment of allowances, reveal that, in most cases, regulation is prompted by a request from one of the utilities in the commission's jurisdiction for a ruling prior to the appearance of any trading activity in that state. But this is not always the case. For instance, the New York State Department of Public Service took the initiative in 1992, prior to any trading activity or requests from utilities for guidelines on allowance treatment, to issue a notice to utilities under its jurisdiction soliciting comments on basic questions regarding the ratemaking treatment of allowances and the role the New York state commission should have in shaping utility emission compliance actions. The casual observation that regulations occur

\textsuperscript{10}There are 21 states that have at least one generating unit affected by Phase 1 emission limitations.
before trading activity in the calendar sense of time also neglects any feedback that particular regulations may have on future allowance trading activity, particularly with respect to the number, volume, or type of allowance transactions that occur.

2.3 EMPIRICAL SPECIFICATION

A reduced form model is used to assess the direct and indirect effects of regulatory and non-regulatory activity on allowance trading behavior. In addition, in reduced form the data can shed light on which of the two hypothetical impacts of PUC regulation on allowance trading behavior dominates. On the one hand, regulation may reduce regulatory uncertainty, decreasing the transaction costs associated with allowance trading activity, thereby increasing allowance trading activity. On the other hand, regulation may provide disincentives to trading, increasing transaction costs. Higher transaction costs decrease allowance trading activity. The data can reveal, in reduced form, which regulatory effect dominates allowance trading behavior.

The empirical method is based on annual observations from 45 states and the District of Columbia (which is referred to as a state for the purposes of this paper) between the years 1993 and 1995. Five states are excluded: Nebraska because it does not have any investor owned utilities, Alaska and Hawaii because they are not affected by Title IV, Idaho because it does not have any fossil fuel fired (SO2 emitting) utilities, and Tennessee because all generating units in Tennessee are both operated and regulated by the Tennessee Valley Authority.

A. Model Specification

Whether a state experiences allowance trading activity by one or more of its electric utilities is modeled as a binary choice.\textsuperscript{11} To address the endogeneity concern, allowance trading activity is

---

\textsuperscript{11} The binary choice model does not differentiate between large volumes of trading activity nor large numbers of trades. As a result, the simultaneous logit model will not capture the possibility that states without allowance regulation have observed only small volumes of trades or small numbers of trades, while those states with regulations observe large volumes of trades or large numbers of trades. This issue is taken up in the next section.
considered in a simultaneous framework by specifying an allowance trading equation and an allowance regulation equation. Assuming a logistic distribution\textsuperscript{12}, the probability that state $s$ experiences allowance trading activity in year $t$ is:

$$P(TRADE_{st}=1) = \exp(\alpha RULING_{st} + \Sigma X_{stk}b_k)/(1+\exp(\alpha RULING_{st} + \Sigma X_{stk}b_k))$$

where $TRADE_{st}=1$ if at least one utility in state $s$ experienced “arms-length” allowance trading activity in year $t$, $RULING_{st}=1$ if there was a regulation on allowance trading activity present in state $s$ in year $t$, and $X_{stk}$ are the $K$ ($k=1,..,K$) non-regulatory factors which influence the decision to engage in allowance trading activity. Assuming a logistic distribution, the probability that state $s$ has a regulation on allowance trading activity in year $t$ is:

$$P(RULING_{st}=1) = \exp(\delta TRADE_{st} + \Sigma Z_{stm}\gamma_m)/(1+\exp(\delta TRADE_{st} + \Sigma Z_{stm}\gamma_m))$$

where $RULING_{st}$ and $TRADE_{st}$ are defined as per above, and $Z_{stm}$ are the $M$ ($m=1,..,M$) non-trading related factors which affect whether a regulation is issued on allowance trading activity.

Following Schmidt and Strauss (1975), a simultaneous logit model is derived. Schmidt and Strauss show that from the above specification it follows that $\delta = \alpha$, and that the appropriate likelihood function to maximize over $\beta$, $\gamma$, and $\alpha$ is:

$$\prod_{i=0}^1 \prod_{j=0}^1 \prod_{i=0}^1 P(RULING_{st} = i, TRADE_{st} = j) \quad (i,j = 0,1)$$

\textsuperscript{12}Although I have no a priori reason to assume a particular probability distribution function, the simultaneous logit model is substantially more tractable than the simultaneous probit model (See Schmidt and Strauss 1975).
where RULING$_{it}$ and TRADE$_{it}$ are as defined above, $\Theta_{i,j} = \{st \mid RULING_{it} = i, TRADE_{it} = j\}$, and $P(RULING_{it} = i, TRADE_{it} = j)$ takes on one of the following functions depending on the value of RULING$_{it}$ and TRADE$_{it}$:

$$
\begin{align*}
\text{P}(\text{RULING}_t=0, \text{TRADE}_t=0) &= 1/\Delta_t \\
\text{P}(\text{RULING}_t=0, \text{TRADE}_t=1) &= \exp(\Sigma X_{stk} B_k)/\Delta_t \\
\text{P}(\text{RULING}_t=1, \text{TRADE}_t=0) &= \exp(\Sigma Z_{stm} \gamma_m)/\Delta_t \\
\text{P}(\text{RULING}_t=1, \text{TRADE}_t=1) &= \exp(\Sigma X_{stk} B_k + \Sigma Z_{stm} \gamma_m + \alpha)/\Delta_t
\end{align*}
$$

where $\Delta_t = 1 + \exp(\Sigma X_{stk} B_k) + \exp(\Sigma Z_{stm} \gamma_m) + \exp(\Sigma X_{stk} B_k + \Sigma Z_{stm} \gamma_m + \alpha)$.

B. Variable Specification and Data

i. Allowance Trading Equation

As discussed above, TRADE$_{it}$=1 if at least one utility in state s experienced "arms-length" allowance trading activity in year t. The primary source for the trading data is the Allowance Transaction System, supplemented with data from reports in Energy Daily, as well as from reports in Clean Air Compliance Review (CACR). Allowance trading activity is expected to depend positively on state PUC regulatory behavior. Two variables assess state PUC behavior. First, a commission more favorable to electric utility shareholder interests (RATING) may be less likely to rule allowance purchases or sales imprudent thereby increasing allowance trading activity. The Merrill Lynch Opinions of Regulation, 1992 – 1995 is used to assess how favorable a PUC commission is to electric utility shareholder interests.$^{13}$ Second, a generic order or guideline (RULING) explicitly issued by a state PUC on how allowances will be treated for ratemaking purposes is expected to minimize utilities' concerns that allowance purchases will be ruled imprudent in a ratemaking case thereby increasing allowance trading activity observed in that state. On the other hand though, regulation may provide disincentives to trading, increasing

---

$^{13}$For a discussion of the use of investment bank ratings as a gauge of the regulatory climate faced by electric utilities see Joskow, Rose and Wolfram 1994 and the literature cited therein.
transaction costs thereby decreasing allowance trading activity. In reduced form, the data can reveal which regulatory effect dominates.

Several non-regulatory factors \((X_k, k=1,...,K)\) are also hypothesized to affect the probability that a state experiences allowance trading activity. First, the aggregate number of allowances the state was allocated by Title IV above or below its annual cost minimizing allowance allocation (ALLOCATION) is expected to have a positive effect on allowance trading activity.\(^{14}\) ALLOCATION is equal to the absolute value of the difference between the state’s aggregate Title IV allowance allocation and the state’s aggregate cost minimizing allowance allocation.\(^{15}\) Utilities in states with more or less allowances than expected to be needed for compliance purposes may have excess allowances to sell or need to purchase additional allowances to, at the very least, meet their cost minimizing allowance needs. The square of the variable ALLOCATION is included in order to capture the effect that states with allowance allocations further from cost minimizing allowance needs will be even more likely to observe allowance trading activity than those states whose allocation differs only slightly from cost minimizing allowance needs.

Second, a greater percentage of a state’s generating capacity subject to Phase 1 emission limitation requirements (UNITs) is expected to have a positive effect on allowance trading behavior in that state. Because the Acid Rain Program was implemented in two Phases, utilities with more Phase 1 affected generating capacity are expected to have assessed their allowance

\(^{14}\)ALLOCATION is not defined simply as the allocation given by Title IV because if each state were allocated its cost minimizing allocation by Congress then no trading should be expected to be observed. The idea behind ALLOCATION is that allowances were not perfectly allocated (see Joskow and Schmalensee 1998), and therefore trading should be expected. Those states that were given “too few” allowances will need to buy allowances and those states that were given “too many” allowances will want to sell allowances.

\(^{15}\)The cost minimizing allowance allocation is the allocation of allowances that minimizes estimated total compliance costs. The cost minimizing allowance allocation is derived from ICF 1990. See also Joskow and Schmalensee 1998.
needs for compliance purposes earlier than Phase 2 affected utilities, and therefore more likely to engage in allowance trading activity.

Allowance trading in previous years by utilities in a particular state (PAST), is expected to have two opposing effects on current trading behavior. First, previous trading behavior is expected to have a positive effect on allowance trading activity through a learning or familiarity effect that acts to minimize perceived risk of allowance trading. On the other hand, previous trading activity is expected to have a negative effect on allowance trading activity to the extent that a utility in a state engaged in trading activity in previous years may have fulfilled its allowance needs and therefore not need to trade allowances in the current year. The variable PAST\textsubscript{st} equals one if utilities in state s traded allowances in a year prior to year t and zero otherwise.

Finally, the immediacy of mandated emission requirements is expected to have a positive effect on the likelihood utilities in a state engage in allowance trading activity. The probability of trading allowances should be smaller in 1993 (THREE) and 1994 (FOUR) relative to 1995 for two reasons. First, planning for the first year of compliance is less immediate in earlier periods. Second, utilities may perceive the risk and uncertainty associated with allowance trading to be greater in earlier years of the program. It is expected that the likelihood of allowance trading activity in 1994 is greater than the likelihood of allowance trading activity in 1993, all else equal.

The specification of the allowance trading equation is thus:

\[
\alpha \text{RULING}_n + \sum x_{ik} B_k = \alpha \text{RULING}_n + \beta_1 \text{RATING}_n + \beta_2 \text{ALLOCATION}_n + \beta_3 \text{ALLOCATION}^2_n + \\
\beta_4 \text{UNITS}_n + \beta_5 \text{PAST}_n + \beta_6 \text{THREE}_n + \beta_7 \text{FOUR}_n + \mu_n
\]
ii. Regulation Equation

As discussed above, $\text{RULING}_{st}=1$ if state $s$ has a regulation on allowance trading activity in year $t$. Allowance trading activity in a state (TRADE) is expected to increase the likelihood that the state will have a regulation on allowance trading activity because it may be more efficient and cost effective to address the issue of allowance trading comprehensively rather than on a continuing case by case basis. Non-trading activity in a state is also expected to affect the decision to regulate allowance trading. A larger percent of the state's generating capacity affected by Phase 1 emission limitations (UNITS) is expected to increase the likelihood that the state will have a regulation on allowance trading since more units subject to emissions restrictions suggests that more utilities may take an active, and immediate, interest in allowance trading. States with elected PUC commissioners (ELECTED) are expected to have a lower likelihood that the state will have a regulation on allowance trading activity because a PUC with elected commissioners is expected to be more observant of public opinion, environmental pressure, or constituent pressure.

The specification of the regulation equation is thus:

$$\delta \text{TRADE}_{st} + \sum_{m} \gamma_m z_{sm} = \delta \text{TRADE}_{st} + \gamma_1 y_1 \text{UNITS}_{st} + \gamma_2 \text{ELECTED}_{st} + \varepsilon_{st}$$

2.4 EMPIRICAL RESULTS

Table 3 presents maximum likelihood estimates from the above likelihood function using data from 46 states inclusive of the District of Columbia in the years 1993 - 1995. The top section of Table 3 presents parameter estimates from the allowance trading specification while the bottom section of Table 3 presents parameter estimates from the regulation specification. A likelihood ratio test was used to test the null hypothesis that $\beta_i = \gamma_j = \alpha = \delta = 0$ for $i=2,...,8$ and $j=2,3$. Denoting $L_0$ as the restricted logit model and $L$ as the unrestricted logit model where all
coefficients are free, \((-2.0)(\ln L_0 - \ln L)\) is distributed as a chi-square with 9 degrees of freedom. The test statistic is 236.26, and the null is rejected at the 1% significance level.

Parameter estimates for the allowance trading specification given in Table 3 are consistent with expectations. With respect to PUC behavior, a PUC favorable to shareholder interests (RATING) has a positive and significant effect on the probability that allowance trading activity is observed in that state. In addition, the presence of a regulation on allowance trading activity (RULING) has a positive and significant effect on the probability that utilities in that state engage in allowance trading activity. A likelihood ratio test was done to assess the null hypothesis that PUC behavior does not affect allowance trading activity. A joint test of the hypothesis that \(\beta_2 = \alpha = 0\) is rejected at the 1% significance level, indicating that PUC behavior positively affects allowance trading activity.

The number of excess allowances issued to a state (_ALLOCATION_) has a negative but insignificant effect on the probability that utilities in that state will engage in allowance trading activity. In addition, the coefficient on the square of the number of excess allowances (\_ALLOCATION^2\_) is positive but insignificant. The percent of generating capacity affected by Phase 1 emissions limitations (UNITS) has a negative but insignificant effect on allowance trading activity. Trading activity observed in a previous period (PAST) has a positive effect on allowance trading activity, but is not statistically significant at the 5% level.

The passage of time has a positive and significant effect on allowance trading activity, suggesting that trading activity is more likely to be observed in 1995 relative to 1993 and 1994. This result suggests that the immediacy of mandated emission requirements had a positive effect on the likelihood utilities in a state engage in allowance trading activity either because planning for the first year of compliance is less immediate in earlier periods or because risk and uncertainty associated with allowance trading is perceived to be greater in earlier years of the
program. A likelihood ratio test was done to test the null hypothesis that the passage of time between 1993 and 1995 (THREE) as well as between 1994 and 1995 (FOUR) has no differential effect on trading activity: $\beta_7=\beta_8$. The null fails to be rejected at the 5% significance level, suggesting that the allowance trading activity is equally less likely in 1993 and 1994 compared to 1995.

Parameter estimates for the regulation specification given in the lower portion of Table 3 are also consistent with expectations. The parameter estimates for UNITS and TRADE are positive and significant suggesting that the percent of generating capacity affected by Phase 1 emissions limitations and observed trading activity increase the probability that the PUC will issue a regulation on allowance trading activity. The variable ELECTED is negative, as expected, but is not statistically significant.

Table 4 presents the parameter estimates from maximum likelihood logit estimation when the allowance trading equation is estimated separately (non-simultaneous) from the regulation equation. As in Table 3, the top section of Table 4 presents estimates from the allowance trading specification, while the bottom section of Table 4 presents estimates from the regulation specification. For the allowance trading specification, a likelihood ratio test of the null hypothesis that $\alpha = \beta_i = 0$ for $i=2,..8$ yields a test statistic of 78.8 with 9 degrees of freedom, clearly rejecting the null hypothesis at the 1% level. For the regulation specification, a likelihood ratio test of the null hypothesis that $\delta = \gamma_i = 0$ for $i=2, 3$ gives a test statistic of 62.30 with 3 degrees of freedom, clearly rejecting the null hypothesis at the 1% level.

By and large the implications from the simultaneous approach compared to those from the single equation approach are very similar. Looking first at the allowance trading equation, the coefficients on RATING, ALLOCATION, ALLOCATION$^2$, THREE, and FOUR are very
similar in magnitude, direction, and significance between the simultaneous approach and non-simultaneous approach. Although the negative sign on the coefficient UNITS in the simultaneous approach is inconsistent with expectations, the variable does not carry any explanatory power. The coefficient on PAST is smaller and insignificant in the simultaneous logit approach. The coefficient on RULING, the endogenous variable, becomes larger and remains significant at the 1% level when one takes the endogeneity into account. Turning to the regulation equation, the coefficients on UNITS and TRADE are very similar in magnitude, direction, and significance between the two estimation approaches. The variable ELECTED loses all explanatory power when one accounts for the endogeneity of trading activity.

2.5 ADDITIONAL EVIDENCE ON THE EFFECT OF STATE PUC REGULATION
A second way state PUC regulations may influence allowance trading behavior is by affecting the extent to which allowance trading activity occurs. For example, both a state with a regulation and a state without a regulation may have utilities engaging in allowance trading activity, but the state with the regulation may have utilities engaging in "more" allowance trading, measured by volume of allowances traded or number of executed transactions, than the state without a regulation on allowance trading activity. Because the logit model considers only the binary trade-no trade decision and ignores any detailed information available on the number and volume of allowance trading activity, the logit model is unable to capture any effect that states without allowance regulations observe only small volumes of trades or small numbers of trades while those states with regulations observe large volumes of trades or large numbers of trades. This section explores the possibility that state PUC regulations effect the extent to which allowance trading activity occurs and finds some evidence that state PUC regulations positively effects the number of executed allowance transactions.
Two equations are estimated in order to assess whether state PUC regulations affect the extent to which allowance trading activity occurs. The first assesses the effect of regulatory activity on the volume of allowance trading. The equation estimated takes the form:

\[
\text{VOLUME}_{st} = \beta_0 + \beta_1 \text{RULING}_{st} + \beta_2 \text{RATING}_{st} + \beta_3 \text{ALLOCATION}_{st} + \beta_4 \text{ALLOCATION}^2_{st} + \beta_5 \text{UNITS}_{st} + \beta_6 \text{PAST}_{st} + \beta_7 \text{THREE}_{st} + \beta_8 \text{FOUR}_{st} + \mu_{st}
\] (1)

where the right hand side variables are as defined in the previous section and \( \text{VOLUME}_{st} \) is equal to the number of allowances traded in state \( s \) in year \( t \). As in the previous section, \( \text{RULING} \) is instrumented for using the variable \( \text{ELECTED} \).

The second equation employs a Poisson model to analyze the effect of state PUC regulations on the number of allowance transactions that occur. The estimated equation takes the form:

\[
\#\text{TRADE}_{st} = e^{\beta_0 + \beta_1 \text{RULING}_{st} + \beta_2 \text{RATING}_{st} + \beta_3 \text{ALLOCATION}_{st} + \beta_4 \text{ALLOCATION}^2_{st} + \beta_5 \text{UNITS}_{st} + \beta_6 \text{PAST}_{st} + \beta_7 \text{THREE}_{st} + \beta_8 \text{FOUR}_{st} + \mu_{st}}
\] (2)

where the right hand side variables are as defined in the previous section and \( \#\text{TRADE}_{st} \) is equal to the number of transactions occurring in state \( s \) in year \( t \).

Table 5 presents the parameter estimates from generalized least squares estimation of equation (1). Parameter estimates in Table 5 suggest that a state PUC regulation on allowance trading activity has no statistically significant discernable effect on the volume of allowances transacted. Results in Table 5 suggest that the volume of allowances transacted is greater in states that were allocated allowances further from cost minimizing allowance needs; the difference is statistically different from zero at the 10% level. This result is consistent with the hypothesis that states
which were allocated many more allowances than needed to meet SO₂ emissions transact a larger volume of allowances because they have additional allowances to sell off and states which were allocated many fewer allowances than needed to meet SO₂ emission needs transact a larger volume of allowances because they need to purchase many more allowances in order to meet SO₂ allowance needs. Finally, parameter estimates in Table 5 suggest that allowance trading in a past period increases the volume of allowances transacted by 88,498 allowances. This difference is statistically different from zero at the 1% level. The positive effect of previous trading activity on volume of allowances traded is consistent with the expectation that a learning or familiarity effect acts to minimize the perceived risk of allowance trading.

Table 6 presents the parameter estimates from a Poisson maximum likelihood estimation of equation (2). Coefficient estimates measure the change in the rate at which allowance trading activity occurs as a result of a one unit increase in the right hand side variable. Coefficient estimates statistically different from one imply a differential effect in the rate at which allowance trading activity occurs. Parameter estimates in Table 6 suggest that the rate at which allowance trading occurs is 2.09 times greater in states with a PUC regulation on allowance trading activity than in states without a PUC regulation on allowance trading activity. This difference is statistically different from one at the 5% level. A state PUC favorable to utility interests increases the rate at which allowance trading occurs by 1.08 times and is significant at the 1% level. A state with a larger percent of generating capacity designated as Phase I affected increases by 1.67 times the rate at which allowance trades occur. This result is significantly different from one at the 1% level. States which have experienced allowance trading activity in previous years have a 3.9 times greater rate of allowance trades activity in the current year. Finally, being in 1993 or 1994 decreases the rate of executed allowance trades by approximately one half. This difference is statistically different from one at the 1% level.
Taken together, Table 5 and 6 suggest that state PUC regulations on allowance trading activity have some effect on the number allowance trades taking place but not on the volume of allowances transacted. All else equal, a state with a regulation may be expected to make more allowance transactions to acquire the same volume of allowances as a state without a regulation on allowance trading activity. One explanation for this phenomenon may be linked to the role that a PUC regulation plays in diminishing the perceived risk associated with allowance trading activity. There are two competing hypotheses. On the one hand, because a PUC regulation may diminish the risk associated with a prudency review, a state with a regulation may be less worried about seeking out the best available transaction and instead purchases smaller bundles of allowances less tailored to their specific needs than a state without a regulation. As a result, utilities in states with a PUC regulation on allowance trading activity are more comfortable incurring transaction costs than utilities in states without a PUC regulation on allowance trading activity. On the other hand, a utility with a PUC regulation may be more comfortable seeking out the best possible bundle of allowances despite the increased transaction costs incurred by purchasing smaller packages of allowances. A utility without a state PUC regulation may perceive multiple transactions and multiple transaction costs as a more risky strategy than a single purchase. Without additional proprietary details on the price and contract terms that allowance bundles were purchased at, it is impossible to disentangle these two competing explanations.16

2.6 IMPLICATIONS AND CONCLUSIONS

When considering allowance trading activity, the question naturally arises whether there would be still more allowance trading activity if state public utility commission rulings were more favorable, such as explicit incentive regulation, or if more state commissions had issued rulings

---

16 The EPA’s Allowance Transaction System, the official record of allowance holdings for compliance purposes and the primary source for allowance trading data, does not require utilities to report prices or terms of trade. The only data required to be reported to the EPA is the quantity of allowances transacted and the names of the two transacting parties.
specific to the ratemaking treatment of allowances. One tempting, but incorrect, argument to make is along these lines:

Fifteen states have issued rulings on allowance trading activity, and arms length allowance trading activity was observed in all fifteen of these states. Thirty-two states have not issued guidelines or generic orders, and no trading was observed in half of these states. Therefore, regulations should be issued in states without regulations in order to encourage allowance trading activity in those states.

There are three inter-related rejoinders to this argument. The first two stem from this analysis and the third rests on basic economic principles. First, it is clear that PUC rulings on the treatment of allowances are not a prerequisite for allowance trading. As this paper discussed, thirty-one states have traded allowances, but only fifteen states have allowance regulations on the books. Put another way, sixteen states have traded allowances without a formal or informal ruling from their state public utility commission.

Second, no state public utility commission that was requested to rule on the ratemaking treatment of allowances from one of its utilities has flatly denied the request. In conversations with the staff of commissions that had yet to concern themselves with the ratemaking treatment of allowances, the reason a ruling had not been issued was typically that the commission had not received any requests for formal or informal guidance from utilities under their jurisdiction. While this is an imperfect, and perhaps biased, measure of the need for regulatory rulings, it does suggest that some utilities are comfortable trading without a formal ruling on how allowances will be treated for ratemaking purposes.

The third and final comment to the assertion that regulation is hindering allowance trading activity is that it is not obvious that more trading activity than that observed to date, in aggregate or in any particular state, should be occurring. That is to say, a large amount of trading, as measuring in terms of volume of allowances traded, number of allowance trades occurring, or
simply as the number of states with utilities trading allowances, is not an indication of how well the market is functioning. It is important to remember that neither the volume of allowances traded, the number of trades, nor the number of states trading allowances reveals much about how well the market is working. More allowance trading, or less for that matter, relative to what has been observed to date, without additional information on cost savings, says very little about the success of the allowance trading program. Only if there are additional cost saving that could have been obtained if additional trading activity had occurred should state regulations, or the lack there of, be suspected to be the cause. Therefore, although the conclusion reached in this analysis is that regulation has had a strong positive effect on allowance trading activity in the early years of the Acid Rain Program, this study should not be interpreted as suggesting that the issuance of guidelines or orders by more state public utility commissions would have lead to more allowance trading.

In summary, it is clear utilities are trading allowances, and that commissions have, by and large, been responsive to utilities' requests for guidance. In addition, the language of the orders encourages, rather than restricts, allowance trading activity. The regulatory rulings that have been issued appear to have had the effect of minimizing the perceived risk of unfavorable rulings on the ratemaking treatment of allowances from trading activity. Until there is evidence suggesting that significant additional cost savings could have been obtained if additional allowance trading activity had occurred in states without regulations or that utilities in states with regulations are still not taking advantage of all cost saving trading opportunities, this analysis suggests that there is little reason to believe that allowance trading activity is impeded by public utility commission regulations.

---

17 An analysis of emissions trading under the U.S. Acid Rain Program by Ellerman et. al. 1997 evaluates compliance costs and allowance market performance and finds that utilities did take advantage of the cost-saving flexibility provided by emissions trading. Whether additional cost saving could have been obtained if additional trading activity had occurred is not addressed.
REFERENCES


U.S. Environmental Protection Agency, Acid Rain Division (EPA, 1995). *Allowance Transactions, TRANSmmd.* Data is available on hard copy from the Acid Rain Division or from the Acid Rain Homepage on the Internet.


### Table 1
States With PUC Guidelines or Generic Orders

<table>
<thead>
<tr>
<th>Guidelines Issued</th>
<th>Generic Order Issued</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>Connecticut*</td>
</tr>
<tr>
<td>Illinois</td>
<td>Georgia</td>
</tr>
<tr>
<td>Maryland</td>
<td>Indiana</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>Iowa</td>
</tr>
<tr>
<td>New York</td>
<td>Mississippi</td>
</tr>
<tr>
<td>Ohio</td>
<td>Missouri</td>
</tr>
<tr>
<td></td>
<td>North Carolina*</td>
</tr>
<tr>
<td></td>
<td>Pennsylvania</td>
</tr>
<tr>
<td></td>
<td>Wisconsin</td>
</tr>
</tbody>
</table>

*States with no Phase I affected units.

Note: To date, the issuance of guidelines or generic orders has been mutually exclusive.

### Table 2
Number of States With PUC Guidelines or Generic Orders, 1993-1995

<table>
<thead>
<tr>
<th>Year</th>
<th>Guidelines: Number of States</th>
<th>Generic Orders: Number of States</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1993</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1994</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>1995</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

Note: To date, the issuance of guidelines or generic orders has been a mutually exclusive event.
Table 3  
Simultaneous Maximum Likelihood Logit Estimates: Parameter Estimates

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Asymptotic Standard Error</th>
<th>Asymptotic Normal Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>RULING</td>
<td>5.372</td>
<td>0.932</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td></td>
<td>-1.545</td>
<td>0.867</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>RATING</td>
<td>0.119</td>
<td>0.062</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>ALLOCATION</td>
<td>-0.0005</td>
<td>0.0004</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>ALLOCATION$^2$</td>
<td>6.3x10^{-8}</td>
<td>4.6x10^{-7}</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>UNITS</td>
<td>-2.445</td>
<td>2.188</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>PAST</td>
<td>0.063</td>
<td>0.653</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>THREE</td>
<td>-1.627</td>
<td>0.642</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>FOUR</td>
<td>-1.226</td>
<td>0.550</td>
</tr>
<tr>
<td>$\delta$</td>
<td>TRADE</td>
<td>5.373</td>
<td>0.932</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td></td>
<td>-7.222</td>
<td>1.118</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>UNITS</td>
<td>4.848</td>
<td>2.639</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>ELECTED</td>
<td>-26.767</td>
<td>8211.4</td>
</tr>
</tbody>
</table>

*significant at 10%, **significant at 5%, ***significant at 1%

N = 138
LogLikelihood, lnL, = -50.03
<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Asymptotic Standard Error</th>
<th>Asymptotic Normal Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>RULING</td>
<td>1.372</td>
<td>0.625</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td></td>
<td>-2.264</td>
<td>0.838</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>RATING</td>
<td>0.244</td>
<td>0.099</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>ALLOCATION</td>
<td>0.0005</td>
<td>0.0004</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>ALLOCATION$^2$</td>
<td>6.7x10^{-7}</td>
<td>3.8x10^{-7}</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>UNITS</td>
<td>3.346</td>
<td>1.865</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>PAST</td>
<td>2.106</td>
<td>0.958</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>THREE</td>
<td>-2.095</td>
<td>0.591</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>FOUR</td>
<td>-1.828</td>
<td>0.652</td>
</tr>
<tr>
<td>$\delta$</td>
<td>TRADE</td>
<td>4.189</td>
<td>0.578</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td></td>
<td>-5.853</td>
<td>0.558</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>UNITS</td>
<td>7.129</td>
<td>1.588</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>ELECTED</td>
<td>-0.0004</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

*significant at 10%, **significant at 5%, ***significant at 1%
N=138.
LogLikelihood for the decision to trade equation (top) is -55.801
LogLikelihood for the decision to issue regulations equation (bottom) is -51.017
Table 5

Generalized Least Squares: Parameter Estimates
Dependent variable = volume of trades (VOLUME)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>9030.9</td>
<td>25998.1</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>7033.4</td>
<td>20538.4</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>161.86</td>
<td>2897.9</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-3.6387</td>
<td>13.161</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>0.0220*</td>
<td>0.01259</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>-36702.6</td>
<td>51553.4</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>88498.7***</td>
<td>20733.6</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>-9315.0</td>
<td>18201.1</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>1459.0</td>
<td>16454.7</td>
</tr>
</tbody>
</table>

*significant at 10%, **significant at 5%, ***significant at 1%
N=138.
Table 5

Poisson Maximum Likelihood Estimation: Parameter Estimates
Dependent variable = number of trades (#TRADE)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>RULING</td>
<td>2.099**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>RATING</td>
<td>1.077***</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>ALLOCATION</td>
<td>0.999</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>ALLOCATION$^2$</td>
<td>1.000</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>UNITS</td>
<td>1.676*</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>PAST</td>
<td>3.909***</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>THREE</td>
<td>0.537***</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>FOUR</td>
<td>0.692***</td>
</tr>
</tbody>
</table>

*statistically different from 1 at 10%, **statistically different from 1 at 5%, ***statistically different from 1 at 1%
N=138.
APPENDIX
PUBLIC UTILITY COMMISSION GUIDELINES AND GENERIC ORDERS

Review of FERC Accounting Guidelines

On March 31, 1993 the Federal Energy Regulatory Commission (FERC) issued Revisions to the Uniform System of Accounts in order to account for allowances. The revisions are not intended to promote or discourage particular ratemaking treatment for allowances, and FERC leaves the revisions open to state PUC variations. Deliberately distinguishing allowances from fuel or financial instruments, allowances are to be classified as Allowance Inventory or Allowances Withheld. Allowances acquired for speculative purposes must be accounted for in "Other Investments". Allowances must be expensed monthly based on each month's SO2 emissions. The FERC revision states that historical cost is the appropriate measure of the accounting value of allowances, but makes clear the distinction that historical cost is not necessarily the best measure for the ratemaking value of allowances and leaves open the option for alternate treatments with respect to ratemaking. All allowances, including those purchased and sold between affiliates, are be accounted for at the purchased price (read historical cost). Other revisions include a call for weighted average cost methods for allowance inventory accounting, and a decision to decline to adopt below the line sharing of gains or losses on the purchase or sale of allowances (except for allowances used for speculative purposes whose gains or losses are kept entirely by the utility).

An Overview of Commission Guidelines and Generic Orders

While all utilities must follow FERC accounting practices, the state public utility treatment of allowances for ratemaking purposes has been varied. Though the majority of states issuing generic ratemaking treatment of allowances are those that are affected most immediately by Phase 1, there are two Phase 2 states, Connecticut and North Carolina, which have taken significant steps to define the treatment of allowances for ratemaking purposes. Ratemaking treatment varies widely among those states which have issued a generic order but tend to require revenues and expenses are passed on 100% to the ratepayers via offsets to fuel costs. The informal guidelines in states which are working on a case by case basis are similar in theme to those developed by states with formal orders. Most state guidelines have requested that allowance expenses and gains to flow to the ratepayers one for one through some type of fuel adjustment clause.

Two Phase 1 states have not concerned themselves with the ratemaking treatment of allowances nor with the revenue generated from the pro rata return of moneys from the EPA annual auctions. Kansas has only one Phase 1 utility, Quindaro, a municipal, which is not subject to the jurisdiction of the Kansas Public Utility Commission. In Tennessee, the Tennessee Valley Authority (TVA) generates electricity for all units in Tennessee. In addition, TVA self-regulates all its electric generating units. As a consequence, TVA simply factors in sulfur dioxide allowances as a cost of doing business and no concern has been raised as to how, or if, revenue from allowance sales (or purchases) should be considered in the ratemaking process. Beyond Kansas and Tennessee, two other Phase 1 states have yet think about the treatment of allowances. Both the Michigan Public Utility Commission and the Minnesota Public Utility Commission are not currently dealing with the issue. Not surprisingly, both Michigan and Minnesota have only one unit each affected in Phase 1.
CONNECTICUT
The Connecticut Public Utility Commission ruling establishes an incentive based approach to the ratemaking treatment of allowances encouraging the sale of allowances. In 1993, Connecticut issued a generic order requiring that 85% of the costs and benefits resulting from allowance transactions on non-bonus allowances be returned to the ratepayers with the other 15% of the benefits retained by the utility. Non-incentive revenue, such as revenue from the pro-rata return of moneys from the EPA advance auctions as well as revenue from the sale of bonus Conservation and Renewable Energy Reserve (CRER) allowances that Connecticut utilities received, were ordered to be returned 100% to the ratepayers. All revenue is returned to the ratepayers by offsetting rate increases due to the costs incurred from state regulations which require utilities to undertake conservation load management.

GEORGIA
In April of 1994, the Georgia Public Service Commission issued a procedural response as a way to answer questions raised by Savannah Electric and Power and Georgia Power Company regarding these utilities plans for compliance with the CAAA. With respect to the accounting and ratemaking treatment of allowances, Georgia Power Company, a Phase I operating company, will include emissions allowances in inventory at cost and will expense them as they are consumed. Utilities are required to flow any allowance gains or losses at market value (rather than at FERC historic cost accounting) through the fuel adjustment clause on an annual bases, and will be treated as a rate reduction for ratemaking purposes. Any allowance profits or losses made on below the line transactions would be kept entirely by the utility. The commission also denied Georgia Power its request to keep fifty percent of all profits on allowances sales from those allowances that Georgia Power acquired from joining the Phase I Extension Pool. The PSC ruled that all profits (or losses) from those allowances would flow to the ratepayers.

INDIANA
Indiana Utility Regulatory Commission has issued a statute which requires up front approval of all allowance purchases and sales, as well as an order which dictates that allowances are the property of ratepayer. As a consequence of a PSI Energy petition for approval of its Environmental Compliance Plan, the Indiana Utility Regulatory Commission ruled that PSI Energy must conform its emissions allowances accounting practices to those laid out by the FERC final rules issued in 1993. Emission allowances used to satisfy off-system loads will be accounted separately from other allowances. A tariff rider will be used to recover allowance costs. Since FERC rules did not explicitly treat banked allowances the Commission decided to make its own rulings on the treatment of banked allowances; the Commission rejected PSI Energy's proposal to add incremental costs to the historic costs of the banked emissions allowances ruling such allowance costs are to be recorded at their acquisitions costs. The Commission goes on to note it will defer issues of carrying charges on allowances purchased specifically for banking to future hearings.

IOWA
In Iowa, all moneys generated from the purchase or sale of allowances flow through to the ratepayer one for one through an energy adjustment clause. In order to recover the cost of purchasing allowances, a rate-regulated utility must file monthly reports with the commission indicating the number and cost of allowances used per month, as well as the number and price of all allowances purchased or sold in that month, and the dollar amount of any gains or losses.

MISSISSIPPI

96
Until December 1995, the Mississippi Public Service Commission had allowed the recovery of all revenues and expenses from allowance transactions to be recovered (or rebated) through an environmental compliance plan. At the request of the Mississippi Power Company, the Commission has revised the means by which revenues and expenses are recovered and has instead incorporated the costs and gains from allowance purchases and sales into the fuel adjustment clause (FAC) and are now recovered in the same manner as other direct fuel expenses. All money's continue to be recovered (or rebated) one for one to ratepayers.

**MISSOURI**
The state of Missouri has a statute which requires any part of utility used to make electricity to be subject to regulation. As a result, the Missouri Public Service Commission issued an order that requires utilities to get prior approval to sell allowances. The order recognizes sales may be decided quickly so in practice the PUC gives blanket permission for all sales. Missouri has had three cases which have raised the issue of allowance expenses and revenues. In each case, the ratemaking treatment of allowances was handled differently. Kansas City Power and Light's allowance sale was approved by the PUC and required KCP&L to defer all revenue until a future rate case. Empire District Electric Company was required to subtract off all annual EPA auction revenue from its fuel costs calculations. The decision in the Union Electric Case was to allow the utility to retain all profits from allowance sales if the profit was less than 11% with any profits in excess of 11% to be split fifty-fifty between the utility and the ratepayers.

**NORTH CAROLINA**
Prompted by a request from Duke Power to accrue a carrying cost on its net investment in allowances purchased in the 1993 EPA auctions (Duke Power purchased 25,000 vintage 2000 allowances for $3,675,000) the North Carolina PUC began discussions on issuing a generic order for allowance treatment. Carolina Power and Light had also notified the North Carolina commission of its purchase of allowances requesting, and receiving, permission to issue a promissory note for the purchase of allowances. In response to the requests, the North Carolina Commission ordered that allowances would be allowed to accrue a carrying charge on those allowances acquired for the purpose of achieving Phase 2 compliance in an analogous way to the accrual charge allowed on cost of work in progress (CWIP). The North Carolina Commission also ordered that no portion of the net investment in allowance inventory would be considered by the Commission for inclusion in the rate base prior to the year 1999. Sales of allowances must be reported to the Commission and the proceeds from such sales are to be used to offset the net investment in allowance inventory.

**PENNSYLVANIA**
In Pennsylvania, the Public Utility Commission's order dictates that any costs associated with pollution control technologies can only be considered a non-revenue producing investment, and recovered through the cost of work in progress (CWIP) clause, if any benefits from the sale of allowances related to that technology are passed on to the utility customers. This clause rings of the allowances that Pennsylvania received as a result of the Phase 1 Extension Bonuses. The Pennsylvania order goes on to require that allowances issued by the EPA be valued at original cost (i.e. zero cost) while purchased allowances will be valued at their full purchase price inclusive of broker fees. Emission allowances are treated as fuel inventory for ratemaking purposes and are recovered through the utility's energy cost rate (ECR). Furthermore, allowances in inventory are to earn a return in the same way as other rate base investments. The order does prohibit two significant actions, both of which Commissioner Wendell Holland dissented with when he announced the formal order. First, the commission's order explicitly prohibits the purchasing of allowance options and futures. Second, the order prohibits cost recovery incentives as part of a utility’s compliance plan such as the retention of the gains from the sale of allowances funded from below the line sources.

**WISCONSIN**
Wisconsin requires utilities to report to the state a filing on what they expect to do over the next twenty years with respect to the use of their allowances including annual streams of allowances expected from the EPA, annual allowance use for compliance, and annual reserve banks. The purpose, says the Public Service Commission of Wisconsin, is to prevent a utility from selling too many allowances and finding itself short in future years. Net revenues from allowance transactions are credited entirely to the ratepayers and are accounted for in materials and supplies in the net investment rate base. A utility is required to notify the Commission after a trade has been made including price, quantity, and the second party(s) involved. The Commission explicitly states no aspect of the sale will be permitted to be confidential unless revealing the second party would cause harm to the ratepayers. Wisconsin makes specific the accounting treatment of allowance transactions between a utility and its holding company or one of its affiliates. Allowance trades between utilities and affiliates in a holding company system require that services or assets provided by a utility to an affiliate be priced at the greater of cost or fair market value.

Public Utility Commissions: Informal Guidelines

ALABAMA
Until allowances come into question during a ratemaking hearing or otherwise become an issue in Alabama, currently only Alabama Power Company units in the lower half of Alabama are affected in Phase 1, the Alabama Public Service Commission rebates revenues and costs from allowance purchases and sales one for one to ratepayers through an Energy Cost Recovery clause. Pro rata auction revenues are returned to ratepayers entirely through a separate accounting clause. All other accounting practices such as historic cost accounting follow straight from the FERC policy statement on the ratemaking treatment of allowances.

FLORIDA
In Florida, the two Phase 1 affected utilities collect allowance expenses and rebate allowance revenues through cost recovery clauses. Gulf Power recovers through an environmental compliance cost recovery clause while Tampa Electric recovers its net revenues through a fuel adjustment clause. Both clauses are adjusted every 6 months, once in the spring and once in the fall. Revenue from the EPA auctions is currently being deferred. The Florida PUC currently allows net revenues generated from below the line allowance transactions to be retained 100 percent by the utility.

ILLINOIS
The Illinois Commerce Commission has issued two orders requiring all gains and losses from allowance purchase and sale be passed on 100 percent to the ratepayers through the fuel adjustment clause. Expenses and revenues are recovered monthly as the allowances are used to match tons of sulfur dioxide emissions. The Commission will value allowances at historical cost. Gains and losses from below the line transactions are kept entirely by the shareholders.

MARYLAND
The Maryland Public Service Commission has issued no final disposition on how allowances will be treated for ratemaking purposes. Potomac Edison has significant amounts of excess allowances due to scrubber installation and are currently in discussion with the Maryland PSC on how to handle the revenues if they are sold. Currently Potomac Edison's allowance revenues and costs are swallowed up in a fuel cost clause. Potomac Electric Power Company (PEPCo) has a tariff (Fuel Rate-Rider) which allows pro rata revenues from the EPA auction to float through the fuel rate being returned dollar for dollar to the ratepayers.
NEW HAMPSHIRE
New Hampshire is poised to deal formally with the treatment of allowances in the next few months. It has been an ongoing issue which is only coming to a formal hearing the first week of December of 1995. Currently all pro rata revenue from the annual EPA auctions is being deferred until after the hearing.

NEW YORK
The New York State Department of Public Service issued an order in 1992 posing 22 questions regarding the ratemaking treatment and policy. The Commission has not acted on the responses to that order. On March 2, 1994, the Commission issued a second order requiring all allowance moneys be deferred until a generic order is issued. The Commission is leaning towards revenue rebate and recovery one for one through a FAC, but there is still some question on that issue. Another strong point of contention yet to be resolved is how to treat allowance swaps and other allowance loans. This contention arises from trades which the Long Island Lighting Company has been involved in.

OHIO
Although Ohio has had four rate cases which have brought up the treatment of allowance revenue, Dayton Power and Light, Centerior Energy, Cincinnati Gas and Electric, and Monongahela Power, the Ohio Public Utilities Commission has yet to issue a generic order. Questions concerning the treatment of allowances have primarily been dealt with case by case because the PUC would like to integrate the treatment of allowances into utilities' Environmental Compliance Plan and as part of their Integrated Resource Plan. The Commission issued a guideline which recognized that carrying charges related to the emissions allowance trading process are legitimate costs of doing business, but the PUC has not specifically addressed this issue. The Ohio guidelines also recognize that all reasonable trading mechanisms such as sales, purchases, futures, leases, and options are legitimate forms of trade and should be evaluated on an equal basis. The guidelines also mandate that each utility submit an allowance trading status report with its EFC audit documenting actual and foregone transactions, allowance holdings, an explanation for why a particular bank level has been maintained, and any adverse experience they have had in trading with other states or encounters with other regulatory authorities. Finally the guidelines suggest that all gains or losses on emission allowance transactions flow through to ratepayers on an energy basis unless the utility created the gain or loss from a below the line transaction in which case all gains may be retained by the utility.
CHAPTER 3

Intertemporal Pricing of Sulfur Dioxide Allowances

3.1 INTRODUCTION

Title IV of the 1990 Clean Air Act Amendments (CAAA)\(^1\) established the first large-scale, long-term environmental program to rely on tradable emissions permits to control pollution. This program was designed to cut acid rain by reducing sulfur dioxide (SO\(_2\)) emissions from electric generating plants to about half their 1980 level, beginning in 1995. Acid rain (or, more properly, acid deposition) occurs when SO\(_2\) and nitrogen oxides react in the atmosphere to form sulfuric and nitric acids, respectively.\(^2\) These acids then fall to earth, sometimes hundreds of miles from their source, in either wet or dry form. The restrictions on SO\(_2\) emissions are applied in two phases. Phase I covers the 263 dirtiest large generating units in the country and requires them, in the aggregate, to reduce their emissions substantially, to about 5.7 million tons per year, during the period 1995 - 1999.\(^3\) Phase II, which begins in 2000, tightens the emissions cap further and extends it to virtually all electric generating units in the continental U.S.

\(^1\) The 1990 CAAA is Public Law 101-549.

\(^2\) The Acid Rain Program also aims to reduce nitrogen oxide (NO\(_x\)) emissions by fifty percent from 1980 levels. The NO\(_x\) aspect permits only limited intra-utility trading among sources.

\(^3\) In fact, an additional 182 generating units that were not otherwise due to become "affected sources" until Phase II became affected in Phase I through special (substitution unit or compensating unit) provisions of a compliance plan for one or more of the original 263 Phase I units. See EPA 1996 for statistics and Joskow and Schmalensee 1998 for summaries of these provisions.
Title IV embraces emissions trading with remarkably few restrictions. The law created *de facto* property rights for emissions, called “allowances” in this program, that can be freely traded and it gives electric utilities complete flexibility in determining how they will comply with their obligations under the law. To allow utilities to take advantage of inter-temporal cost savings, Title IV allows utilities to bank unused allowances for use in future years. Allowances can be traded nationally, though the environmental problem being addressed is regional.\(^4\) No review or prior approval of trades is necessary. The purchase and holding of allowances is not restricted to the utilities for whom these permits would become a necessary input for the generation of electricity. All sources receiving allowance allocations as well as third parties, such as brokers and individuals, are free to buy or sell allowances with any other party. Neither the frequency nor the mechanisms for trading allowances are limited. Finally, Section 416(d)(2) of the 1990 CAAA withheld a small fraction (2.8%) of the allowances to be issued to utilities and mandated that they be sold at an annual revenue neutral Environmental Protection Agency (EPA) administered auction. The vast majority of allowance trading has involved bilateral private trades between utilities that own electric generators or between those utilities and third parties, although some allowances have been traded through the annual set of auctions that the EPA is required to hold.\(^5\)

Economists have long argued that, in theory, the tradable permit approach to pollution control should involve significantly lower costs than the traditional command-and-control approach of

\(^4\) The concern that regional patterns of emissions reductions would result in “hot spots” have not been realized (EPA 1996).

\(^5\) Joskow, Schmalensee, and Bailey 1998.
specifying source-specific standards. This argument, of course, rests on the assumption that the market for permits is reasonably efficient. The limited experience with emissions trading prior to 1990 was not particularly encouraging to this end (Hahn 1989, Hahn and Hester 1989), and there was considerable doubt whether this feature of Title IV would meet with any greater success than had earlier experiments with emissions trading.

Joskow, Schmalensee, and Bailey (1998), denoted herein as JSB, rely on “spot” allowance prices reported by private market making organizations and clearing prices in the “spot” EPA auctions as well as bidding behavior in the annual EPA auctions to demonstrate that the market for SO₂ allowances had become reasonably efficient by mid-1994, a full year and a half before utilities had to relinquish to the EPA allowances to cover their emissions in the first year of the program. The close alignment of prices quoted from several independent sources in conjunction with the flattening of the bidders’ offer curves in the annual EPA auctions strongly suggest the emergence of a competitive market for SO₂ allowances.

This paper provides additional evidence toward the proposition that the SO₂ allowance market has become reasonably efficient. Because SO₂ abatement strategies involve investment and contracting decisions that affect emissions for many years, a forward market for allowances should be expected to emerge if utilities take advantage of opportunities for inter-temporal cost savings made possible by the banking provisions of Title IV. Moreover, because Title IV allows unused allowances to be

---

6 On the economics of the tradable permit approach, see Tietenberg 1985.

7 Smaller scale programs employing tradable permits to phase-out leaded gasoline and CFCs have been met with somewhat more success (NERA 1994).
carried forward for use in future years, a predictable term structure should be observed between current and future allowances. In fact, an active inter-temporal market for SO$_2$ allowances has emerged. In this paper, the forward market for SO$_2$ allowances is examined and found to be reasonably efficient, consistent with the hypothesis put forth by JSB.$^8$

The organization of the remainder of the paper is as follows. The next section outlines Title IV of the 1990 CAAA, the associated allowance trading system, and the role of forward market transactions. Section 3 examines, in theory, the intertemporal pattern of allowance prices expected to be observed in the SO$_2$ allowance market. Section 4 considers data from the forward market for allowances and the term structure of allowance prices. Section 5 presents conclusions.

**3.2 TITLE IV AND INTERTEMPORAL ALLOWANCE TRADING**

The basic approach to emissions control embodied in Title IV is simple: An aggregate annual cap on national SO$_2$ emissions defines the number of emission allowances available for allocation to electricity generating units each year. An emissions allowance is the right to emit one ton of SO$_2$ into the atmosphere. To emit SO$_2$ legally during a given year, an affected unit (one subject to Title IV’s SO$_2$ constraint) must have enough allowances that are good for use in that year to cover all its SO$_2$ emissions. Title IV also requires each affected unit to have a continuous emissions monitoring system on each exhaust stack to measure actual SO$_2$ emissions and to report those emissions to the EPA. At the end of each year, each source must have deposited enough allowances in an account

---

$^8$Options and other derivative instruments have also emerged in the SO$_2$ allowance market in recent years. In addition, an active swap market has emerged. A utility which needs allowances today but expects to have excess allowances in 2000 will often find it tax-efficient to “swap” allowances with another entity rather than buy allowances today and sell allowances in, say, 2000. See Bartels 1997.
maintained for it by the EPA to cover all of its recorded emissions or be subject to significant financial (and legal) penalties.

Title IV specifies the initial allocation of SO₂ allowances. Allowances are given to existing electric generating units and those under construction according to fairly complicated rules that are discussed in detail by Joskow and Schmalensee (1998). For purposes of this paper it suffices to note that essentially all of the allowances available to cover SO₂ emissions were allocated "free" to incumbent sources.

Each allowance specifies a particular year, its "vintage", in which it is first available to be used to cover SO₂ emissions. An allowance can also be "banked" and used in any future year. Thus, for instance, a "1996 vintage" allowance can be used to cover emissions in 1996 or held for use in any later year, but it cannot be used to cover 1995 emissions. Most importantly, all allowances are fully tradable. That is, a source that has been allocated allowances is free to sell them to any other source, including to a third party such as brokers and individuals. Moreover, an affected source can buy allowances to cover its present emissions or its future emissions from any type of trading partner. There are no limitations on how often parties can trade allowances or on the trading mechanisms that buyers and sellers may use to trade them. Finally, Title IV created a set of small revenue-neutral allowance auctions in order to "jump-start" the market. The auctions, which are administered by the EPA, occur only once a year and cover only two or three vintages at a time. Each year, roughly 2.8% of the allowances that have been allocated to utilities are held back and auctioned in annual "spot" and "seven-year advance" auctions. In addition, in 1994 through 1997 a "six-year advance" auction was also held because allowances available in the "Direct Sales
Reserve” provision of the law were not sold. Allowances sold in the seven (six) year advance auction are first usable seven (six) years after the auction. For example, in the 1993 seven-year advance auction, vintage 2000 allowances were sold; in the 1994 seven-year advance auction, vintage 2001 allowances were sold, and so on. The revenues from all these auctions are returned to the sources in proportion to their share of the allowances that were held back for the auctions.

At the very least, a forward market for SO₂ allowances should be expected to develop in order for participants to price the six and seven year advance auction allowances sold by the EPA since auction participants can purchase the spot allowances sold in the EPA auction and carry them forward for six or seven years or alternatively simply purchase the six or seven year advance allowances. Moreover though, a robust private market for forward allowances should be expected to develop since compliance decisions by electric utilities are often long term investment decisions frequently requiring streams of allowances involving many consecutive vintages. In addition, utilities with excess vintage 1995 allowances to sell, for example, typically also have excess allowances of several adjacent vintages to sell. Similarly, utilities needing to purchase vintage 2000 allowances, for example, typically also need to purchase several adjacent vintages as well. Moreover, utilities responding to current and future allowance needs require the flexibility to buy or sell allowances at any time not just in the year the allowances are needed for compliance purposes.

Table 1 reports information on the number of spot⁹ and forward allowances that appear to have been traded in arms-length transactions through December 1997. Following JSB, Table 1 is

---

⁹ A 1995 spot allowances is a 1995 allowance. Because allowances can be banked, a 1996 spot allowance is a 1995 or a 1996 allowance and a 1997 spot allowance is a 1995, 1996 or 1997 allowance.
constructed using data from the EPA’s Allowance Tracking System in order to create a conservative lower bound estimate of the volume of arms-length trading during various time periods between 1992, when the first trades took place, and December 31 1997. Arms length transactions include both private market transactions as well as EPA auction transactions. The details of how estimates of private market trades have been constructed are available in JSB. For a variety of reasons discussed in JSB, the recorded data in the EPA’s allowance tracking system may underestimate actual commercial transactions to some extent.

Table 1 demonstrates that there has in fact been considerable arms-length trading in forward (non-spot) allowances. By the end of December 1997, approximately 9.2 million non-spot vintage allowances had been traded in arms length transactions, slightly less than half of the total number of allowances of all vintages, 19.5 million, transacted in the same period. It is also evident from Table 1 that there was little forward vintage allowance transactions prior to the start of 1995. About 8.8 of the 9.2 million forward allowances that were traded up through the end of December 1997 were traded after December 1993 and about 7.8 of the 9.2 million forward allowances that were traded were traded after December 1994. In addition, Table 1 indicates that a larger number of forward vintage allowances denominated in the years 2000-2005 were traded compared to forward vintage allowances denominated in remaining Phase 1 years and compared to vintage year allowances denominated in the years 2006-2025. There are two reasonable explanations for why a larger number of forward allowances denominated in vintages between 2000 and 2005 were traded compared to other forward vintage allowances. First, the six and seven year advance auction allowances that have been sold in the annual EPA auctions all fall into the 2000-2005 category of
vintage allowances.\textsuperscript{10} The second explanation is that the year 2000 is the first year the more stringent Phase 2 emissions limitations become effective. As a result, vintage 2000 allowances are a focal point for utilities’ compliance planning purposes. Utilities appear to transact in spot allowances to meet Phase 1 allowance needs and transact in early Phase 2 allowances to plan for Phase 2 allowance needs. For example, there has been considerable swap activity between spot allowances and vintage 2000 allowances (Bartels 1997). Moreover, because Phase 2 extends to virtually all electric generating units in the continental U.S., significantly more utilities are affected by Title IV starting in the year 2000 and must plan accordingly for compliance purposes, including allowance needs, in the year 2000.

The volume of arms-length allowances that have been traded in the forward market has given rise to readily available forward prices from intermediaries and other third party brokers.\textsuperscript{11} The next section lays out the term structure expected to be observed in a competitive and reasonably efficient allowance market.

3.3 THEORETICAL FRAMEWORK

In a world of certainty without transaction costs, if there is banking across all relevant periods, then arbitrage between current and expected future compliance costs and between allowances of differing vintages will cause the immediate settlement prices for allowances of different vintages

\textsuperscript{10} From 1993 through 1996, 100,000 allowances were offered for sale in the seven-year advance auction (vintages 2000 to 2003 respectively). From 1994 through 1997, because of the operation of the “Direct Sales Reserve” provision of the law, 25,000 allowances were offered for sale in the six-year advance auction (2000-2003). Because the EPA closed down the “Direct Sales Reserve” beginning in 1997, the 25,000 allowances involved were included in the 1997 seven-year advance auction, which accordingly involves 125,000 vintage 2004 allowances.

\textsuperscript{11} Of course, readily available forward prices is also likely to give rise to more forward allowance trading activity.
to be equal. In all periods, an individual holder of allowances chooses a level of SO₂ abatement such that the current marginal cost of abating in that period equals the current price of a spot market allowance \((MC_t = P_t)\). Across any two periods with banking, the individual abates so that the discounted marginal cost of abatement is equal \((MC_t = MC_r)\). Since the discounted marginal cost of abatement is equal across periods and current marginal cost of abatement is equal to the current allowance price in every period, the immediate settlement prices for allowances of differing vintages are equal \((P_t = P_r)\). That is to say, allowances of differing vintages should sell for the same price today. For example, the immediate settlement price for a vintage 1996 allowance in 1996 should equal the immediate settlement price of a vintage 1997 allowance in 1996. This also implies that the spot price of an allowance should grow at the rate of interest when banking occurs between period \(t\) and period \(t+1\), \textit{ceteris paribus}.

The allowance market, though, is marked by uncertainty and positive transaction costs. One source of uncertainty is the cost of compliance. As a result, a firm may desire to revise its SO₂ emissions should abatement costs turn out to be different than expected. Second, actual emissions of SO₂ are stochastic even when abatement costs are certain, and thus, a firm's need for SO₂ allowances is stochastic. For these two reasons, SO₂ emissions in any given compliance year cannot be known with absolute certainty.

The allowance market is also marked by positive transaction costs, though the magnitude of these costs have declined over time. One cost is the commission fee paid to the market maker or broker for rendering his services. Commissions per allowance averaged $1.75 in mid-1994, $1.00 in late 1995, $0.75 in September, 1996, and $0.50 in early 1997. The latter figure was less than one
percent of the prevailing spot price. A second type of transaction cost is the search costs incurred from deciding which broker to use and the search costs incurred if the firm decides instead to act as its own broker. A third transaction cost is the cost of having to negotiate a contract and the accompanying terms of sale. Search costs may decline over time as firms become more experienced with executing transactions in the allowance market.

Because SO$_2$ emissions in any given year are uncertain and it is costly to purchase and sell allowances, a firm benefits from holding a stock of allowances on hand to buffer itself against unexpectedly high SO$_2$ emissions. The benefit that accrues from holding a stock of allowances on hand, called a convenience yield$^{12}$, is the transaction cost saved from not having to make additional transactions and/or undo the transaction just done.$^{13}$ To the extent that there is uncertainty and positive transaction costs, allowances of different vintages will not sell for the same price today.

A model incorporating convenience yields generates weak backwardation in the term structure of allowance prices. The return on the risk-free portfolio of holding an allowance and shorting a forward $T$ periods is the sum of the forward price ($F_T$) less the spot market price today ($P_1$) plus the convenience yield ($CY_T$):

$^{12}$ Wright and Williams 1991.

$^{13}$ For the case of undoing the transaction just done, consider the example of the firm which sells off all unused 1995 allowances in 1995 (technically, it would be at the end of the “true-up” period on 31 January 1996), and then finds itself with unexpectedly high emissions in the following year - due to unexpectedly high costs or just the stochastic nature of SO$_2$ emissions in general. The firm must now return to its broker (or find a trading partner) and repurchase allowances, incurring the same set of transaction costs a second time.
\[(F_T - P_1) + CY_T.\]  

(1)

Since the portfolio is risk-free, its return must equal the risk-free return. The risk-free return is the return from selling the allowance on the spot market today and subsequently investing that money in an asset with a certain return, such as a Treasury bill, until period \(T\):

\[r_T P_1,\]  

(2)

where \(r_T\) is the rate of interest between today and period \(T\). Equating (1) and (2) and rearranging gives:

\[\frac{F_T}{(1 + r_T)} + \frac{CY_T}{(1 + r_T)} = P_1\]  

(3)

Rewriting equation (3) in terms of the discounted forward price \(f_T\), the discounted convenience yield \(cy_T\), and the spot price gives:

\[f_T + cy_T = P_1\]  

(4)

For any positive convenience yield then, allowance prices will be weakly backwarded: \(f_T < P_1\).

For example, a positive convenience yield will cause the immediate settlement price for a vintage
1996 allowance transacted in 1996 to be greater than the immediate settlement price of a vintage 1997 allowance transacted in 1996. An immediate settlement price is the price paid today for the purchase of an allowance today.

As the convenience yield declines toward zero, because uncertainty in the allowance market resolves or transaction costs decline toward zero, allowances of different vintages will sell for the same price today: $f_T = P_1$. For example, if the convenience yield equals zero, then the immediate settlement price for a vintage 1996 allowance transacted in 1996 should equal the immediate settlement price of a vintage 1997 allowance transacted in 1996.

3.4 EMPIRICAL EVIDENCE

Monthly immediate settlement prices for current and future vintage allowances published by Cantor Fitzgerald Environmental Brokerage Services, an intermediary in the SO₂ allowance market, can be used to assess the term structure of allowance prices. The immediate settlement (year $t$) price for a current vintage allowance (vintage year $t$) is a standard spot market price. Because a vintage year $t+T$ allowance is not first usable until year $t+T$, immediate settlement transactions in year $t$ involving allowances with vintage years greater than the current year can be thought of as a $T$ year forward contract with the date of settlement being the vintage year.\textsuperscript{14} For

\textsuperscript{14} In a standard forward market transaction a price, valued in \textit{year $T$ dollars}, quantity, and settlement date (\textit{year $T$}) are agreed upon today by the two parties; transfer of the money and asset occurs on the settlement date (\textit{year $T$}). Because a vintage year $T$ allowances is not first usable until year $T$, an immediate settlement agreement for a vintage year $T$ allowance can be thought of as a $T$ year forward contract where the transaction price is a discounted futures price. In the interpretation of an immediate settlement transactions involving allowances with vintage years greater than the current year as a forward transaction, a price, valued in \textit{today's dollars}, quantity, and settlement date (\textit{today}) are agreed upon today by the two parties; transfer of the money and asset occurs on the settlement date (\textit{today}).
example, the immediate settlement transaction of a vintage 1997 allowance in the current year, say January 1996, can be thought of as a 1 year forward contract with a date of settlement in 1997. The immediate settlement price in year t for a vintage year t+T allowance is a discounted future price.  

Figure 1 shows that the immediate settlement prices of different vintage allowances are roughly equal over periods for which banking is reasonably expected to occur. Each horizontal line on the Figure represents a single month of immediate settlement data for successive allowance vintages. In Figure 1, horizontal lines marked by diamonds indicate that the monthly set of immediate settlement price data is from 1995, triangles indicate the monthly set of price data is from 1996, and squares represent monthly price data from 1997. Isolating a single horizontal line, discounted prices would be equal if the line is flat, that is, if the immediate settlement price for a 1996 allowance equals the immediate settlement price of a 1997 allowance, equals the immediate settlement price of a 1998 allowance, et cetera. This is in fact just about the pattern observed for immediate settlement price data in 1996 and 1997.

The two immediate settlement price series from 1995 are more obviously tilted backward (i.e. the immediate settlement price of a future vintage year allowance is below the immediate settlement

---

15 Similar types of forward transactions occur in the market for vintage wines. Wine is purchased today at a price denominated in today’s dollars and the transfer of the wine from the seller to the buyer occurs today, but the buyer is not able to use the wine until some date in the future (i.e. when the wine matures or ripens).

16 Estimates from Ellerman and Montero 1996 and those from the CEEPR questionnaire put the year in which the allowance bank in Phase I will be exhausted between 2004 and 2005. The CEEPR questionnaire is a survey distributed to utilities in the U.S. which are affected by Phase I of Title IV of the Clean Air Act of 1990. Responses came from 138 Phase I units, a response rate of 30%.
price of a nearer vintage year allowance). Though less obvious at the scale of Figure 1, 1996 and 1997 immediate settlement prices also exhibit some degree of backwardation.

Table 2 provides numerical evidence on the term structure of current and future vintage allowances. The Table shows the difference between the immediate settlement price of a spot allowance and the immediate settlement price of a forward allowance for several forward vintage years. As noted above, if the private market were frictionless and perfectly competitive, the difference between the immediate settlement price of a spot allowance and the immediate settlement price of a forward allowance would be zero. The smaller the differences shown in Table 2, the more closely the term structure of current and future vintage allowances conforms to the competitive, frictionless market ideal.

Table 2 illustrates several features of the term structure of current and future vintage year allowances. First, the magnitude of the convenience yield is quite small absolutely and quite small relative to the immediate settlement price of a spot allowance. For instance, on average in 1996, the immediate settlement price of a spot allowance was greater than the immediate settlement price of a one year forward allowance by $0.81, or 1% of the spot market price. Second, the magnitude of the convenience yield increases as the vintage of the forward allowance increases. That is, the convenience yield attached to holding a 1995 vintage allowance one year (as opposed to buying a 1996 vintage allowance in 1995) is smaller than the convenience yield attached to holding a 1995 vintage allowance two years (as opposed to buying

---

17 Prices from every months of the year are not illustrated on Figure 1 so not to clutter the graph. Including data from all months of the year does not qualitatively change Figure 1.
a 1997 vintage allowance in 1995). This is reasonable since there is greater uncertainty attached to buying a 1997 allowance in 1995 as opposed to buying a 1996 allowance in 1995.

Finally, Table 2 demonstrates that the magnitude of the convenience yield, and therefore the degree of backwardation, has fallen over time. For example, in 1995 the average convenience yield attached to carrying a spot allowance for two years (as opposed to buying a 1997 allowance in 1995) was $5.36, in 1996 the average convenience yield was $1.04, and by 1997 this average convenience yield was $0.56. The decline in the convenience yield indicates a flattening of the term structure of allowances prices. The term structure of the monthly immediate settlement prices after mid-1996 is considerably flatter than the monthly immediate settlement prices in earlier months, approximating the term structure of allowance prices expected in a world of certainty with no transaction costs.

There are several explanations for the decline in the convenience yield over time, all which are consistent with the efficient market hypothesis put forth by JSB. First, the flattening of the term structure may reflect a reduction in transaction costs. As discussed above, transaction costs have fallen over time resulting in a decline in the penalty a firm must pay if the firm decides to revise its allowance needs in the future. A decline in transaction costs results in a decline in the magnitude of the convenience yield. In addition, the flattening of the term structure of allowances is consistent with increased market liquidity and the resolution of uncertainty about the value of current and future vintage allowances. The decline in uncertainty decreases the magnitude of the convenience yield. Finally, as compliance years pass, affected utilities are likely to have become more experienced in estimating expected SO₂ emissions and measuring
actual SO₂ emissions. As a result, uncertainty over deviations from expected emissions is likely to have diminished over time. As uncertainty resolves, the benefit from holding a stock of allowances on hand to buffer against unexpectedly high SO₂ emissions decreases thereby decreasing the magnitude of the convenience yield. More likely than any single factor, it seems clear that the flattening of the term structure reflects a combination of all four factors.

3.5 CONCLUDING OBSERVATIONS

It was clear to its proponents that the success of Title IV’s innovative tradable allowance program for reducing sulfur dioxide emissions, like that of any tradable permit program, depended critically on the emergence of an efficient private market for rights to emit. The empirical analysis of the forward market for sulfur dioxide allowances indicates that a relatively efficient forward market developed in a few years’ time, by at least early-1996. This result is consistent with the hypothesis put forth by JSB.

When Title IV was first proposed, many observers argued that the conservative electric utility industry would never encourage the development of an efficient market for emission rights. This paper in conjunction with JSB has shown that these critics were wrong. Indeed, if the electric utility industry can so rapidly be engendered to participate in a market for SO₂ allowances, it is reasonable to expect that efficient markets for emission rights will be the norm as long as trading is permitted with few restrictions on when and how those rights can be traded.
REFERENCES


Table 1

Allowances Sold in EPA Auction and Private Market*, By Vintage

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Through 1993</td>
<td>250,430</td>
<td>100,060</td>
<td>151,003</td>
<td>51,106</td>
</tr>
<tr>
<td>Total</td>
<td>10,357,850</td>
<td>2,216,573</td>
<td>4,493,250</td>
<td>2,490,671</td>
</tr>
</tbody>
</table>

* The number of allowances sold in the private market includes inter-utility trades, trades between utilities and third parties, and trades between two non-utility parties. This number excludes intra-utility trades (including intra-holding company trades), reallocations, and options to trade which have not been exercised.

<table>
<thead>
<tr>
<th>Month</th>
<th>Difference between spot and 1 year forward</th>
<th>Difference between spot and 2 year forward</th>
<th>Difference between spot and 4 year forward</th>
<th>Difference between spot and 7 year forward</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 1995</td>
<td>0.25</td>
<td>5.00</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Oct. 1995</td>
<td>1.75</td>
<td>5.75</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1995 Average</td>
<td>1.00</td>
<td>5.36</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Jan. 1996</td>
<td>1.50</td>
<td>1.75</td>
<td>2.25</td>
<td>3.75</td>
</tr>
<tr>
<td>Feb. 1996</td>
<td>1.08</td>
<td>1.29</td>
<td>1.78</td>
<td>3.04</td>
</tr>
<tr>
<td>March 1996</td>
<td>1.00</td>
<td>1.25</td>
<td>1.60</td>
<td>2.75</td>
</tr>
<tr>
<td>July 1996</td>
<td>0.20</td>
<td>0.45</td>
<td>1.00</td>
<td>2.35</td>
</tr>
<tr>
<td>Aug. 1996</td>
<td>0.30</td>
<td>0.50</td>
<td>1.10</td>
<td>2.25</td>
</tr>
<tr>
<td>1996 Average</td>
<td>0.81</td>
<td>1.04</td>
<td>1.54</td>
<td>2.82</td>
</tr>
<tr>
<td>April 1997</td>
<td>0.10</td>
<td>0.35</td>
<td>1.25</td>
<td>2.80</td>
</tr>
<tr>
<td>July 1997</td>
<td>0.85</td>
<td>1.00</td>
<td>1.75</td>
<td>2.95</td>
</tr>
<tr>
<td>Aug. 1997</td>
<td>0.39</td>
<td>0.61</td>
<td>1.27</td>
<td>2.55</td>
</tr>
<tr>
<td>Nov. 1997</td>
<td>0.03</td>
<td>0.29</td>
<td>1.05</td>
<td>2.53</td>
</tr>
<tr>
<td>1997 Average</td>
<td>0.34</td>
<td>0.56</td>
<td>1.33</td>
<td>2.71</td>
</tr>
</tbody>
</table>